

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 24, 2004, 06:16:38 ; Search time 3026 Seconds
(without alignments)
9754.331 Million cell updates/sec

Title: US-10-068-486a-1

Perfect score: 681
Sequence: 1 ctgaagctctctccacgag.....agagatagctgagtcgac 681

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 2167151695 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*
1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_secs:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_secs:*
28: em_un:*
29: em_vl:*
30: em_hcg_hum:*
31: em_hcg_inv:*
32: em_hcg_other:*
33: em_hcg_mus:*
34: em_hcg_pln:*
35: em_hcg_rod:*
36: em_hcg_mam:*
37: em_hcg_vrt:*
38: em_sy:*
39: em_hcg_hum:*
40: em_hcg_mus:*
41: em_hcg_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	679.4	99.8	681	6	AX276968
2	634.6	93.2	1207	9	HSAN01497
3	634.6	93.2	1253	6	AR400627
4	634.6	93.2	1253	6	AX013131
5	634.6	93.2	1253	6	BD206312
6	634.6	93.2	1546	6	BD234896
7	634.6	93.2	1546	6	BD063242
8	633	93.0	1332	9	BC001055
9	633	93.0	1787	6	AX098214
10	630.8	92.6	636	9	BT006829
11	630.8	92.6	636	12	BT007520
12	622	91.3	1554	6	AX382115
13	579.2	85.1	828	9	AF093823
14	513.8	75.4	894	10	BC008104
15	513.8	75.4	1235	10	BC050007
16	513.8	75.4	158198	2	AC012013
17	512.8	75.3	636	6	E31608
18	512.8	75.3	636	10	AF087825
19	460.2	67.6	754	10	RN0011832
20	384	56.4	384	6	AX351329
21	372.6	54.7	413	6	AR341389
22	372.6	54.7	413	6	BD209583
26	283.8	41.7	241856	2	AC125714
27	283.8	41.7	247766	2	AC111932
28	283	41.6	952	6	E31590
29	283	41.6	952	10	AF072127
30	281.4	41.3	1137	10	BC002003
31	278.4	40.9	1356	5	BC048771
32	278.4	40.9	1356	5	AB072910
33	276.4	40.6	3275	10	BC061992
34	273.4	40.1	893	10	AF195500
35	270.2	39.7	1237	9	AF134160
36	270.2	39.7	1281	9	BC012471
37	270.2	39.7	3400	6	BD18436
38	269.8	39.6	1256	6	BD107883
39	269.8	39.6	1256	6	BD205639
40	268.8	39.5	170227	9	AC092474
41	268.8	39.5	173374	9	AC108882
42	268.6	39.4	3443	6	AR404003
43	268.6	39.4	3443	6	AX342328
44	268.6	39.4	3443	6	BD073876
45	268.6	39.4	3443	9	AF101051

ALIGNMENTS

RESULT 1
AX276968
LOCUS
DEFINITION Sequence 1 from Patent WO0110382.
ACCESSION AX276968
VERSION AX276968.1 GI:16548651
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
1 Nacht, M.
AUTHORS
TITL
JOURNAL
Compositions and methods related to claudin-7
Patent: WO 0110382-A 1 15-FEB-2001;

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 24, 2004, 05:14:11 ; Search time 399 Seconds

(without alignments)
7250.684 Million cell updates/sec

Title: US-10-068-486a-1

Perfect score: 681
1 ctggaagctctgtccaggg.....aggagatgtgtgagtcgac 681

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database:

N_Geneseq_29Jan04:*
1: geneseq1980s:*
2: geneseq1980s:*
3: geneseq2000s:*
4: geneseq2001as:*
5: geneseq2001bs:*
6: geneseq2002s:*
7: geneseq2003as:*
8: geneseq2003bs:*
9: geneseq2003cs:*
10: geneseq2004a:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	679.4	99.8	681	5	AAF60261	Aaf60261 Human Cla
2	634.6	93.2	1014	3	AAC95568	Aac95568 Human sec
3	634.6	93.2	1070	7	ABT31935	Abt31935 Human bre
4	634.6	93.2	1207	8	ADB70321	Adb70321 Claudin-7
5	634.6	93.2	1253	2	AAZ42051	Aaz42051 Human end
6	634.6	93.2	1546	3	AAZ47929	Aaz47929 Human apo
7	634.6	93.2	1583	4	AAH34904	Aah34904 Human col
8	634.6	93.2	1711	2	AAV33617	Aav33617 Human sec
9	634.2	93.0	1504	6	ABQ54998	Abq54998 Human ova
10	633	93.0	1787	5	AAE98719	Aae98719 Human lat
11	622	91.3	1554	6	ABZ35511	Abz35511 Human gen
12	622	91.3	1554	6	ABZ35511	Abz35511 Human gen
13	622	91.3	2252	6	ABZ35548	Abz35548 Human gen
14	519.6	75.3	1686	3	AACT7763	Aac77763 Human can
15	512.8	75.3	636	3	AAZ89154	Aaz89154 Murine cl
16	460.2	75.3	754	7	ABT41698	Abt41698 Toxicity
17	384	56.4	384	6	ABK27639	Abk27639 Human col
18	372.6	54.7	413	3	AAZ61660	Aaz61660 cDNA enco
19	372.6	54.7	413	4	AAC99595	Aac99595 Skin cell
20	372.6	54.7	413	6	ABL34745	Abi34745 Human cDN
21	370.2	54.4	388	6	ABL37118	Abi37118 Human col
22	283	41.6	952	3	AAZ89136	Aaz89136 Murine cl
23	273.4	40.1	893	9	ADB58645	Adb58645 Toxicity-

24	273.4	40.1	893	9	ADB53320	Adb53320 Primary r
25	270.2	39.7	1195	6	ABQ54868	Abq54868 Human ova
26	270.2	39.7	3400	3	AAZ60459	Aaz60459 cDNA enco
27	270.2	39.7	3437	2	ADB47448	Adb47448 Human cDN
28	269.8	39.6	1256	2	AAZ06263	Aaz06263 Human sec
29	269.8	39.6	1256	2	AAZ06263	Aaz06263 Human sec
30	269.8	39.6	1256	9	ADE11645	Ade11645 Human sec
31	268.6	39.5	3433	7	ABX63240	Abx63240 Human cDN
32	268.6	39.4	3443	2	ABX19461	Abx19461 Human sen
33	268.6	39.4	3443	6	ABK15497	Abk15497 Human sen
34	268.6	39.4	3443	7	ABX76250	Abx76250 Lung can
35	268.6	39.4	3443	7	ABX76426	Abx76426 Lung can
36	268.6	39.4	3443	7	ACF12843	Act12843 Human cer
37	268.6	39.4	3445	7	ADA19282	Ada19282 Human ins
38	268.6	39.4	3483	2	AAZ19956	Aaz19956 Human tan
39	268.4	39.4	2747	2	AAZ34118	Aaz34118 Human PRO
40	268.4	39.4	2747	3	AAZ78513	Aaz78513 Human PRO
41	268.4	39.4	2747	4	AAZ45975	Aaz45975 Human DNA
42	268.4	39.4	2747	7	ABX78578	Abx78578 Human PRO
43	268.4	39.4	2747	7	ACA75550	Act75550 Novel hum
44	268.4	39.4	2747	7	ACA71030	Act71030 Human sec
45	268.4	39.4	2747	7	ACC87558	Acc87558 Human sec

ALIGNMENTS

RESULT 1
ID AAF60261 standard; cDNA, 681 BP.
XX
AC AAF60261;
XX
DT 02-MAY-2001 (first entry)
XX
DE Human Claudin-7 cDNA.
XX
KW Human; Claudin-7; cytosolic; angiogenesis inhibition;
XX endothelial cell proliferation; cancer; tumour growth inhibitor; ss.
XX
OS Homo sapiens.
XX
PN WO200110382-A2.
XX
PD 15-FEB-2001.
XX
PF 07-AUG-2000; 2000WO-US021474.
XX
PR 06-AUG-1999; 99US-0147752P.
XX
PA (GENZ) GENZYME CORP.
XX
PI Nacht M;
XX
DR WPI; 2001-202700/20.
XX
P-PSDB; AAB69999.
XX
PT Modulating angiogenesis and/or endothelial cell proliferation and
XX reducing tumor growth in a mammal, involves administering a composition
XX comprising Claudin-7.
XX
PS Example 1; Fig 1; 26pp; English.
XX
CC The present sequence encodes Claudin-7. The invention relates to a method
XX for modulating angiogenesis and/or endothelial cell proliferation in a
XX mammal. The method involves administering a composition comprising at
XX least a fragment, biological equivalent or derivative of Claudin-7, which
XX modulates angiogenesis and/or endothelial cell proliferation, and which
XX reduces tumour growth
XX
SQ Sequence 681 BP; 119 A; 188 C; 211 G; 163 T; 0 U; 0 Other;
XX
Query Match 99.8%; Score 679.4; DB 5; Length 681;

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 24, 2004, 09:44:43 ; Search time 87 Seconds

(without alignments)
4343.926 Million cell updates/sec

Title: US-10-068-486A-1

Perfect score: 681

Sequence: 1 ctggaagtcgtctccaggg.....aggagatgtctgagtcgac 681

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	634.6	93.2	1253	4 US-09-673-395A-71	Sequence 71, Appl
2	372.6	54.7	413	3 US-09-188-930-55	Sequence 55, Appl
3	372.6	54.7	413	4 US-09-312-283C-55	Sequence 55, Appl
4	268.6	39.4	3443	4 US-09-886-683A-3	Sequence 3, Appl
5	268.6	39.4	3443	4 US-09-130-491-3	Sequence 3, Appl
6	258.4	37.9	413	4 US-09-621-976-360	Sequence 360, App
7	210.8	31.0	316	4 US-09-702-705-976	Sequence 976, App
8	210.8	31.0	316	4 US-09-736-457-976	Sequence 976, App
9	210.8	31.0	316	4 US-09-614-124B-976	Sequence 976, App
10	182.4	26.8	1665	4 US-09-671-325-976	Sequence 976, App
11	182.4	26.8	1665	4 US-09-300-958A-22	Sequence 22, Appl
12	181.4	26.6	1705	4 US-09-205-258-106	Sequence 106, Appl
13	180.6	25.7	1722	4 US-09-489-847-93	Sequence 93, Appl
14	174.8	25.5	1708	4 US-09-620-312D-986	Sequence 986, App
15	159	23.3	1380	4 US-09-489-847-125	Sequence 125, App
16	145	21.3	1524	4 US-09-663-600A-15	Sequence 45, Appl
17	144.2	21.2	1400	4 US-09-663-600A-139	Sequence 139, Appl
18	126.6	18.6	363	4 US-08-966-316-8	Sequence 8, Appl
19	114	16.7	441	4 US-09-404-879A-104	Sequence 104, App
20	114	16.7	441	4 US-09-338-933-104	Sequence 104, App
21	114	16.7	441	4 US-09-215-681-104	Sequence 104, App
22	114	16.7	441	4 US-08-216-003A-104	Sequence 104, App
23	112.8	16.6	1165	4 US-09-603-552-2	Sequence 2, Appl
24	101.4	14.9	615	4 US-09-621-976-1835	Sequence 1835, App
25	94.4	13.9	1353	4 US-09-724-864-32	Sequence 32, Appl
26	77.4	11.4	2243	3 US-09-084-079-4	Sequence 4, Appl
27	75.6	11.1	427	4 US-09-603-552-8	Sequence 8, Appl

28	63.4	9.3	1801	1 US-08-557-917A-1	Sequence 1, Appl
29	63.4	9.3	1801	3 US-09-084-153-1	Sequence 1, Appl
30	63.4	9.3	1801	3 US-08-084-079-1	Sequence 1, Appl
31	46.2	6.8	253	4 US-09-603-552-6	Sequence 6, Appl
32	45.6	6.7	541	4 US-09-603-552-7	Sequence 7, Appl
33	44.4	6.5	1212	4 US-09-252-991A-14382	Sequence 14382, A
34	44.4	6.5	1512	4 US-09-252-991A-14321	Sequence 14321, A
35	43.6	6.4	208	4 US-09-603-552-10	Sequence 10, Appl
36	43.2	6.3	394	4 US-09-300-958A-21	Sequence 21, Appl
37	42.2	6.2	1288	1 US-08-440-856A-9	Sequence 9, Appl
38	40.6	6.0	1278	4 US-09-252-991A-738	Sequence 738, App
39	40.6	6.0	1359	4 US-08-252-991A-705	Sequence 705, App
40	40.6	6.0	1587	4 US-09-252-991A-676	Sequence 676, App
41	40.2	5.9	495	4 US-09-252-991A-14290	Sequence 14290, A
42	40	5.9	43280	2 US-08-804-227C-1	Sequence 1, Appl
43	39.4	5.8	1010	3 US-08-858-003-29	Sequence 29, Appl
44	38.4	5.8	1010	3 US-09-078-166-29	Sequence 29, Appl
45	39.4	5.8	1010	3 US-08-997-467-29	Sequence 29, Appl

ALIGNMENTS

RESULT 1									
US-09-673-395A-71									
Sequence 71, Application US/09673395A									
Patent No. 6620923									
GENERAL INFORMATION:									
APPLICANT: SPECHT, THOMAS									
APPLICANT: HINZMANN, BERNARD									
APPLICANT: SCHMITT, ARMIN									
APPLICANT: PILARSKI, CHRISTIAN									
APPLICANT: DAHL, EDGAR									
APPLICANT: ROSENTHAL, ANDRE									
TITLE OF INVENTION: HUMAN NUCLEIC ACID SEQUENCES FROM UTERUS TUMOR TISSUE									
FILE REFERENCE: ALBRE-12									
CURRENT APPLICATION NUMBER: US/09/673.395A									
NUMBER OF SEQ ID NOS: 637									
SOFTWARE: PatentIn Ver. 2.1									
SEQ ID NO 71									
LENGTH: 1253									
TYPE: DNA									
ORGANISM: Homo sapiens									
US-09-673-395A-71									
Query Match									
Best Local Similarity 93.2%; Score 634.6; DB 4; Length 1253;									
Matches 637; Conservative 0; Mismatches 4; Indels 0; Gaps 0;									
QY	35	CGGAGATGCGCAATTTGGGGCTGAGTTGCTGGGGCTTCTCCATGAGCCCTGCTGGGCTGGG	94						
DB	341	CGGAATGCGCAATTTGGGGCTGAGTTGCTGGGGCTTCTCCATGAGCCCTGCTGGGCTGGG	400						
QY	95	TGGGCTGCTGGGCTGCAACCGCCATCCCGAGTGGCAGATGAGCTCTATCGGGGTGACA	154						
DB	401	TGGGCTGCTGGGCTGCAACCGCCATCCCGAGTGGCAGATGAGCTCTATCGGGGTGACA	460						
QY	155	ACATCATGAGCGCCAGGCGCATGTACAAAGGGGCTGTGATGAGTGGTGCACGAGAGCA	214						
DB	461	ACATCATGAGCGCCAGGCGCATGTACAAAGGGGCTGTGATGAGTGGTGCACGAGAGCA	520						
QY	215	CGGGATGATGAGCTGCAAAATGTAGCACTGTGTCTGCGCCCTGCGGGCTTGGCAGG	274						
DB	521	CGGGATGATGAGCTGCAAAATGTAGCACTGTGTCTGCGCCCTGCGGGCTTGGCAGG	580						
QY	275	CCACTCGAGCCCTATGTTGTTCTCCCTGGTGGTGGGCTTCCTCGGCATGTTGTGGCA	334						
DB	561	CCACTCGAGCCCTATGTTGTTCTCCCTGGTGGTGGGCTTCCTCGGCATGTTGTGGCA	640						
QY	335	CGATGGGATGAGTGCACGCGCTGTGGGGAGACGACAAAGTGAAGAGCCCGTATAG	394						
DB	641	CGATGGGATGAGTGCACGCGCTGTGGGGAGACGACAAAGTGAAGAGCCCGTATAG	700						

QY 395 CCATGGGTGGAGGCAATATTTTCATCGTGGCAAGCTCTGGCCCTGGTGGTGGTCTCTCT 454
DB 701 CCATGGGTGGAGGCAATATTTTCATCGTGGCAAGCTCTGGCCCTGGTGGTGGTCTCTCT 760
QY 455 GGTATGGCCATCAGATTGTGACAGACTTTTAAACCTTTGATCCCTAACCAACATTAAGT 514
DB 761 GGTATGGCCATCAGATTGTGACAGACTTTTAAACCTTTGATCCCTAACCAACATTAAGT 820
QY 515 ATGAGTTGGCCCTGGCCATCTTTATTTGGTGGGAGGGTCTGGCCCTGATCTATCTGGGAG 574
DB 821 ATGAGTTGGCCCTGGCCATCTTTATTTGGTGGGAGGGTCTGGCCCTGATCTATCTGGGAG 880
QY 575 GTGACATGCT 634
DB 881 GTGACATGCT 940
QY 635 GCTCTTACCCTTAAGTCCAACTCTTCCAGAGAGTATGTGA 675
DB 941 GCTCTTACCCTTAAGTCCAACTCTTCCAGAGAGTATGTGA 981

RESULT 2

US-09-188-930-55
Sequence 55, Application US/09188930A

Patent No. 6150502

GENERAL INFORMATION:

APPLICANT: Watson, James D.

APPLICANT: Strachan, Lorna

APPLICANT: Sleeman, Matthew

APPLICANT: Onrust, Rene

APPLICANT: Murison, James Greg

TITLE OF INVENTION: Compositions Isolated From Skin Cells

FILE REFERENCE: 11000.1011c1

CURRENT APPLICATION NUMBER: US/09/188,930A

CURRENT FILING DATE: 1998-11-09

NUMBER OF SEQ ID NOS: 348

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 55

LENGTH: 413

TYPE: DNA

ORGANISM: Human

US-09-188-930-55

Query Match 54.7%; Score 372.6; DB 3; Length 413;

Best Local Similarity 98.7%; Pred. No. 9.8e-89;

Matches 386; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

QY 235 ATGTACGACTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 294
DB 18 AGGTACGACTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 77
QY 295 GTCTCCCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 354
DB 78 GTCTCCCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 137
QY 355 CGCTGTGGGGAGAGACCAAGTGAAGAGCCCGTATAGCCATGGTGGAGGCAATAT 414
DB 138 CGCTGTGGGGAGAGACCAAGTGAAGAGCCCGTATAGCCATGGTGGAGGCAATAT 197
QY 415 TTGATCGTGGGAGGCTGGTGGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 474
DB 198 TTGATCGTGGGAGGCTGGTGGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 257
QY 475 ACAGACTTTTATTAACCTTTGATCCCTAACCAATTAAGTATGAGTTGGCCCTGGCATC 534
DB 258 ACAGACTTTTATTAACCTTTGATCCCTAACCAATTAAGTATGAGTTGGCCCTGGCATC 317
QY 535 TTTATTTGGTGGGAGGCTGGTGGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 594
DB 318 TTTATTTGGTGGGAGGCTGGTGGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 376

QY 595 TGTCTGGGATGATGAGCAAGGCTGGTACC 625
DB 377 TGTCTGGGATGATGAGCAAGGCTGGTACC 407

RESULT 3

US-09-112-283C-55
Sequence 55, Application US/09112283C

Patent No. 6573095

GENERAL INFORMATION:

APPLICANT: Watson, James D.

APPLICANT: Strachan, Lorna

APPLICANT: Sleeman, Matthew

APPLICANT: Onrust, Rene

APPLICANT: Murison, James G.

APPLICANT: Kumble, Krishanand D.

TITLE OF INVENTION: Compositions Isolated from Skin Cells

FILE REFERENCE: 11000.1011c2

CURRENT APPLICATION NUMBER: US/09/312,283C

CURRENT FILING DATE: 1999-05-14

NUMBER OF SEQ ID NOS: 425

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 55

LENGTH: 413

TYPE: DNA

ORGANISM: Mouse

US-09-312-283C-55

Query Match 54.7%; Score 372.6; DB 4; Length 413;

Best Local Similarity 98.7%; Pred. No. 9.8e-89;

Matches 386; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

QY 235 ATGTACGACTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 294
DB 18 AGGTACGACTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 77
QY 295 GTCTCCCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 354
DB 78 GTCTCCCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 137
QY 355 CGCTGTGGGGAGAGACCAAGTGAAGAGCCCGTATAGCCATGGTGGAGGCAATAT 414
DB 138 CGCTGTGGGGAGAGACCAAGTGAAGAGCCCGTATAGCCATGGTGGAGGCAATAT 197
QY 415 TTGATCGTGGGAGGCTGGTGGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 474
DB 198 TTGATCGTGGGAGGCTGGTGGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 257
QY 475 ACAGACTTTTATTAACCTTTGATCCCTAACCAATTAAGTATGAGTTGGCCCTGGCATC 534
DB 258 ACAGACTTTTATTAACCTTTGATCCCTAACCAATTAAGTATGAGTTGGCCCTGGCATC 317
QY 535 TTTATTTGGTGGGAGGCTGGTGGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 594
DB 318 TTTATTTGGTGGGAGGCTGGTGGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 376

RESULT 4

US-09-886-683A-3
Sequence 3, Application US/0988683A

Patent No. 6627439

GENERAL INFORMATION:

APPLICANT: Hoevel, Thorsten

APPLICANT: Koch, Stefan

APPLICANT: Kubbies, Manfred

APPLICANT: Mundigl, Olaf

APPLICANT: Rueger, Petra

TITLE OF INVENTION: Antibodies against SEMP1 (p23)

QY	12	CCGAGGAGCAGATGAGCTCCATATGGAGGGTGAACAATCATATGACGGCCAGGCCATGTAC	18
Db	121	CCCCATGTGCAGATGAGCTCTTATGCGGGTGAACAATCATATACCGGCCAGCCATGTAC	18
QY	181	AAGGGGCTGTGATGATGACATCTGCTACGCAAGACACGGGGATGATGAGCTGCAGAAATGTAC	24
Db	181	AAGGGGCTGTGATGATGACATGCTGCACGCAAGACACGGGGATGATGAGCTGCAGAAATGTAC	24
QY	241	GACTCGATGCTGCAGCCGTGTCGGGCGCCTTGCAGAGCCATCGAGCCCTTAATGATGATGCC	30
Db	241	GACTCGATGCTGCAGCCGTGTCGGGCGCCTTGCAGAGCCATCGAGCCCTTAATGATGATGCC	30
QY	301	CTGTGCTGGGCTTCTCTGACATGTTTGTGTGSCCAAGATGAGCATGAGTGAACGCGCTGT	36
Db	301	CTGTGCTGGGCTTCTCTGACATGTTTGTGTGSCCAAGATGAGCATGAGTGAACGCGCTGT	36
QY	361	GGGGGAGACGACAAAGAGAAAGAGAGCCCTATAGCATGGGTGAGAGCATTAATTTTATC	42
Db	361	GGGGGAGACGACAAAGAGAAAGAGAGCCCTATAGCATGGGTGAGAGCATTAATTTTATC	42
QY	421	GTGGCAGGCTTTTGCCCGCTTGTGTGATGCTCTGTATGSCCATCAGATTGTACAAGAC	48
Db	421	GTGGCAGGCTTTTGCCCGCTTGTGTGATGCTCTGTATGSCCATCAGATTGTACAAGAC	48
QY	481	TTTTATTAACCTTGTGATNCCCTACCAACATTAATGATGATTTGGCCCTGGCCACTTATTT	54
Db	481	TTTTATTAACCTTGTGATNCCCTACCAACATTAATGATGATTTGGCCCTGGCCACTTATTT	54
QY	541	GGCTGGGCGAGGCTCTGCTCCTAGTCATCTCTGGAGGTGCACTGCTTCTCTGTTCTGTCTCT	60
Db	541	GGCTGGGCGAGGCTCTGCTCCTAGTCATCTCTGGAGGTGCACTGCTTCTCTGTTCTGTCTCT	60
QY	601	GGGAAATGAGAGCAAGCTGGGTATCCTGTGACCCCGCTCTTAACCTTAATGTCCACTCTTCC	66
Db	601	GGGAAATGAGAGCAAGCTGGGTATCCTGTGACCCCGCTCTTAACCTTAATGTCCACTCTTCC	66
QY	661	AAGAGATATGTGTAGTTCAC	681
Db	661	AAGAGATATGTGTAGTTCAC	681

```

RESULT 2
US-10-176-847-83
? Sequence 83, Application US/10176847
? Publication No. US2003068635A1
? GENERAL INFORMATION:
? APPLICANT: Veldy, Petter Ole
? TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR
? TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF BREAST
? ERLI REFERENCE: MRI-039
? CURRENT APPLICATION NUMBER: US/10/176,847
? NUMBER OF SEQ ID NOS: 112
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ ID NO 83
? LENGTH: 1070
? TYPE: DNA
? ORGANISM: Homo sapiens
? US-10-176-847-83

```

Query Match	93.2%;	Score 634.6;	DB 14;	Length 1070;
Best Local Similarity	99.4%;	Pred. No. 2..2e-176;		
Matches 637; Conservative	0;	Mismatches 4;	Indels 0;	Gaps 0

QY	DB	QY	DB
35	CGCGAATGCGCAATGGGGCTTGGCAATCTCTGGCTTCCATGGCCCTGGCTGGCTGG	95	TGGGTCTGCTGGGCTGACCGCCATCCCGCAGTGGCAATAGAGCTCTTAATCGGGGTACA
131	CGGAATGCGCAATCGGGCTTCAATTCTGGGCTTCCATGGCCCTGGCTGGGCTGGG	191	TGGGTCTGCTGGGCTGACCGCCATCCCGCAGTGGCAATAGAGCTCTTAATCGGGGTACA

[illegible]

```

RESULT 3
US-10-236-031B-13
: Sequence 13, Application US/10236631B
: Publication No. US20030219760A1
: GENERAL INFORMATION:
:   APPLICANT: Gordon, Gavin J.
:   APPLICANT: Jensen, Roderrick V.
:   APPLICANT: Gullans, Steven R.
:   APPLICANT: Bueno, Raphael
:   TITLE OF INVENTION: Diagnostic and Prognostic Tests
:   FILE REFERENCE: B00801/70265 (JRV/JAV)
:   CURRENT APPLICATION NUMBER: US/10/236,031B
:   CURRENT FILING DATE: 2002-09-05
:   PRIOR APPLICATION NUMBER: US 60/317,389
:   PRIOR FILING DATE: 2001-09-05
:   PRIOR APPLICATION NUMBER: US 60/407,431
:   PRIOR FILING DATE: 2002-08-30
:   NUMBER OF SEQ ID NOS: 102
:   SOFTWARE: PatentIn version 3.1
:   SEQ ID NO 13
:   LENGTH: 1207
:   TYPE: DNA
:   ORGANISM: Homo sapiens
US-10-236-031B-13

```

Query Match	93.2%	Score 634.6	DB 15	Length 1207
Best Local Similarity	99.4%	Pred. No. 2.3e-176		
Matches 637; Conservative	0	Mismatches 4	Indels 0	Gaps 0

QY 35 CCGCAGTGGCCATTGGGACCCTGACGTGTGGGCTTCCATGAGCCCTGCTGGGCTGGG 94

Db 330 CGAAATGGCCAAATTGGGCGCTGACAGTGTGGGCTTCCATGAGCCCTGCTGGGCTGGG 389

QY 95 TGGGTCTGGTGGCCCTGACACCGCCATCCGACAGTGGCAGATGAGCTCCTATGCGGATGACA 154

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 24, 2004, 12:18:43 (Search time 4096 Seconds
(without alignments)
5850.036 Million cell updates/sec

Title: US-10-068-486a-1
Perfect score: 681
Sequence: 1 cctgaagcttcgtccaggg.....aggagatcgtgagtcgac 681

Scoring table: IDENTITY_NTC
Gapop 10.0 , Gapext 1.0

Searched: 37577330 seqs, 17593059518 residues

Total number of hits satisfying chosen parameters: 75154660

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Pending Patents NA Main:*

- 1: /cgn2_6/prodata/2/pna/US006_COMB.seq:*
- 2: /cgn2_6/prodata/2/pna/US07_COMB.seq:*
- 3: /cgn2_6/prodata/2/pna/US08_COMB.seq:*
- 4: /cgn2_6/prodata/2/pna/US08_COMB.seq:*
- 5: /cgn2_6/prodata/2/pna/US08_COMB.seq:*
- 6: /cgn2_6/prodata/2/pna/US08_COMB.seq:*
- 7: /cgn2_6/prodata/2/pna/US08_COMB.seq:*
- 8: /cgn2_6/prodata/2/pna/US08_COMB.seq:*
- 9: /cgn2_6/prodata/2/pna/US08_COMB.seq:*
- 10: /cgn2_6/prodata/2/pna/US08_COMB.seq:*
- 11: /cgn2_6/prodata/2/pna/US08_COMB.seq:*
- 12: /cgn2_6/prodata/2/pna/US08_COMB.seq:*
- 13: /cgn2_6/prodata/2/pna/US08_COMB.seq:*
- 14: /cgn2_6/prodata/2/pna/US08_COMB.seq:*
- 15: /cgn2_6/prodata/2/pna/US08_COMB.seq:*
- 16: /cgn2_6/prodata/2/pna/US08_COMB.seq:*
- 17: /cgn2_6/prodata/2/pna/US08_COMB.seq:*
- 18: /cgn2_6/prodata/2/pna/US08_COMB.seq:*
- 19: /cgn2_6/prodata/2/pna/US08_COMB.seq:*
- 20: /cgn2_6/prodata/2/pna/US08_COMB.seq:*
- 21: /cgn2_6/prodata/2/pna/US08_COMB.seq:*
- 22: /cgn2_6/prodata/2/pna/US08_COMB.seq:*
- 23: /cgn2_6/prodata/2/pna/US08_COMB.seq:*
- 24: /cgn2_6/prodata/2/pna/US08_COMB.seq:*
- 25: /cgn2_6/prodata/2/pna/US08_COMB.seq:*
- 26: /cgn2_6/prodata/2/pna/US08_COMB.seq:*
- 27: /cgn2_6/prodata/2/pna/US08_COMB.seq:*
- 28: /cgn2_6/prodata/2/pna/US08_COMB.seq:*
- 29: /cgn2_6/prodata/2/pna/US08_COMB.seq:*
- 30: /cgn2_6/prodata/2/pna/US08_COMB.seq:*
- 31: /cgn2_6/prodata/2/pna/US08_COMB.seq:*
- 32: /cgn2_6/prodata/2/pna/US08_COMB.seq:*
- 33: /cgn2_6/prodata/2/pna/US08_COMB.seq:*
- 34: /cgn2_6/prodata/2/pna/US08_COMB.seq:*
- 35: /cgn2_6/prodata/2/pna/US08_COMB.seq:*
- 36: /cgn2_6/prodata/2/pna/US08_COMB.seq:*
- 37: /cgn2_6/prodata/2/pna/US08_COMB.seq:*
- 38: /cgn2_6/prodata/2/pna/US08_COMB.seq:*
- 39: /cgn2_6/prodata/2/pna/US08_COMB.seq:*
- 40: /cgn2_6/prodata/2/pna/US08_COMB.seq:*
- 41: /cgn2_6/prodata/2/pna/US08_COMB.seq:*
- 42: /cgn2_6/prodata/2/pna/US08_COMB.seq:*
- 43: /cgn2_6/prodata/2/pna/US08_COMB.seq:*

- 44: /cgn2_6/prodata/2/pna/US100B_COMB.seq:*
- 45: /cgn2_6/prodata/2/pna/US101A_COMB.seq:*
- 46: /cgn2_6/prodata/2/pna/US101B_COMB.seq:*
- 47: /cgn2_6/prodata/2/pna/US102A_COMB.seq:*
- 48: /cgn2_6/prodata/2/pna/US102B_COMB.seq:*
- 49: /cgn2_6/prodata/2/pna/US103A_COMB.seq:*
- 50: /cgn2_6/prodata/2/pna/US103B_COMB.seq:*
- 51: /cgn2_6/prodata/2/pna/US104A_COMB.seq:*
- 52: /cgn2_6/prodata/2/pna/US104B_COMB.seq:*
- 53: /cgn2_6/prodata/2/pna/US106_COMB.seq:*
- 54: /cgn2_6/prodata/2/pna/US107A_COMB.seq:*
- 55: /cgn2_6/prodata/2/pna/US107B_COMB.seq:*
- 56: /cgn2_6/prodata/2/pna/US6000_COMB.seq:*
- 57: /cgn2_6/prodata/2/pna/US6001_COMB.seq:*
- 58: /cgn2_6/prodata/2/pna/US6002_COMB.seq:*
- 59: /cgn2_6/prodata/2/pna/US6003_COMB.seq:*
- 60: /cgn2_6/prodata/2/pna/US6004_COMB.seq:*
- 61: /cgn2_6/prodata/2/pna/US6005_COMB.seq:*
- 62: /cgn2_6/prodata/2/pna/US6006_COMB.seq:*
- 63: /cgn2_6/prodata/2/pna/US6007_COMB.seq:*
- 64: /cgn2_6/prodata/2/pna/US6008_COMB.seq:*
- 65: /cgn2_6/prodata/2/pna/US6009_COMB.seq:*
- 66: /cgn2_6/prodata/2/pna/US6010_COMB.seq:*
- 67: /cgn2_6/prodata/2/pna/US6011_COMB.seq:*
- 68: /cgn2_6/prodata/2/pna/US6012_COMB.seq:*
- 69: /cgn2_6/prodata/2/pna/US6013_COMB.seq:*
- 70: /cgn2_6/prodata/2/pna/US6014_COMB.seq:*
- 71: /cgn2_6/prodata/2/pna/US6015_COMB.seq:*
- 72: /cgn2_6/prodata/2/pna/US6016_COMB.seq:*
- 73: /cgn2_6/prodata/2/pna/US6017_COMB.seq:*
- 74: /cgn2_6/prodata/2/pna/US6018_COMB.seq:*
- 75: /cgn2_6/prodata/2/pna/US6019_COMB.seq:*
- 76: /cgn2_6/prodata/2/pna/US6020_COMB.seq:*
- 77: /cgn2_6/prodata/2/pna/US6021_COMB.seq:*
- 78: /cgn2_6/prodata/2/pna/US6022_COMB.seq:*
- 79: /cgn2_6/prodata/2/pna/US6023_COMB.seq:*
- 80: /cgn2_6/prodata/2/pna/US6023A_COMB.seq:*
- 81: /cgn2_6/prodata/2/pna/US6024_COMB.seq:*
- 82: /cgn2_6/prodata/2/pna/US6025_COMB.seq:*
- 83: /cgn2_6/prodata/2/pna/US6026_COMB.seq:*
- 84: /cgn2_6/prodata/2/pna/US6027_COMB.seq:*
- 85: /cgn2_6/prodata/2/pna/US6028_COMB.seq:*
- 86: /cgn2_6/prodata/2/pna/US6029_COMB.seq:*
- 87: /cgn2_6/prodata/2/pna/US6030_COMB.seq:*
- 88: /cgn2_6/prodata/2/pna/US6031_COMB.seq:*
- 89: /cgn2_6/prodata/2/pna/US6032_COMB.seq:*
- 90: /cgn2_6/prodata/2/pna/US6033_COMB.seq:*
- 91: /cgn2_6/prodata/2/pna/US6034_COMB.seq:*
- 92: /cgn2_6/prodata/2/pna/US6035_COMB.seq:*
- 93: /cgn2_6/prodata/2/pna/US6036_COMB.seq:*
- 94: /cgn2_6/prodata/2/pna/US6037_COMB.seq:*
- 95: /cgn2_6/prodata/2/pna/US6038_COMB.seq:*
- 96: /cgn2_6/prodata/2/pna/US6039_COMB.seq:*
- 97: /cgn2_6/prodata/2/pna/US6040_COMB.seq:*
- 98: /cgn2_6/prodata/2/pna/US6041_COMB.seq:*
- 99: /cgn2_6/prodata/2/pna/US6042_COMB.seq:*
- 100: /cgn2_6/prodata/2/pna/US6043_COMB.seq:*
- 101: /cgn2_6/prodata/2/pna/US6044_COMB.seq:*
- 102: /cgn2_6/prodata/2/pna/US6045_COMB.seq:*
- 103: /cgn2_6/prodata/2/pna/US6046_COMB.seq:*
- 104: /cgn2_6/prodata/2/pna/US6047_COMB.seq:*
- 105: /cgn2_6/prodata/2/pna/US6048_COMB.seq:*
- 106: /cgn2_6/prodata/2/pna/US6049_COMB.seq:*
- 107: /cgn2_6/prodata/2/pna/US6050_COMB.seq:*
- 108: /cgn2_6/prodata/2/pna/US6051_COMB.seq:*
- 109: /cgn2_6/prodata/2/pna/US6052_COMB.seq:*
- 110: /cgn2_6/prodata/2/pna/US6053_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	ORGANISM: Homo sapiens	
;	FEATURE:	
;	NAME/KEY: CDS	
;	LOCATION: (40)..(672)	
;	OTHER INFORMATION:	
US-10-068-486A-1		
Query Match	100.0%; Score 681; DB 44; Length 681;	
Best Local Similarity	100.0%; Pred. No. 2,9e-149;	
Matches 681; Conservative 0; Mismatches 0; Indels 0; Gaps 0		
Qy	1 CTGGAAGTTCTGTTCACGGGCCCCCTCGCGCGAGTCCGAGATGCGCAATTGGGCGCTTGACG	60
Db	1 CTGGAAGTTCTGTTCACGGGCCCCCTCGCGCGAGTCCGAGATGCGCAATTGGGCGCTTGACG	60
Qy	61 TTGCTAGGCTTCCATAGGCGCTGTGAGGCTGGGGGTGTGGGGCTGTGACCGGCATC	120
Db	61 TTGCTAGGCTTCCATAGGCGCTGTGAGGCTGGGGGTGTGGGGCTGTGACCGGCATC	120
Qy	121 CCGCAGTGGCAATGAGCTCTATGCGGGTACAACATCATCAGGCCACAGCCCATGTAC	180
Db	121 CCGCAGTGGCAATGAGCTCTATGCGGGTACAACATCATCAGGCCACAGCCCATGTAC	180
Qy	181 AAGGGCTGTGATGATGACCTGGCTCAGACGAGACAGCGGGAGATGAGACTGTGAAAATTAC	240
Db	181 AAGGGCTGTGATGATGACCTGGCTCAGACGAGACAGCGGGAGATGAGACTGTGAAAATTAC	240
Qy	241 GACTCGATCTGCGCTGTGCGGCGCTTGGAGGCACTCAGGCCCTAATGATGATGCC	300
Db	241 GACTCGATCTGCGCTGTGCGGCGCTTGGAGGCACTCAGGCCCTAATGATGATGCC	300
Qy	301 CTGGTGTGAGGCTTCTGCGCATGTTGTGGCCAGATGGGCATGAAGTGCACCGGCT	360
Db	301 CTGGTGTGAGGCTTCTGCGCATGTTGTGGCCAGATGGGCATGAAGTGCACCGGCT	360
Qy	361 GGGGAGACGACAAAGTAGAGAGGCGCGTATGACATGAGGTGGAGGCATTAATTTTCATC	420
Db	361 GGGGAGACGACAAAGTAGAGAGGCGCGTATGACATGAGGTGGAGGCATTAATTTTCATC	420
Qy	421 GTGGCAGGCTTTCGCGCTTGGTACTGTTCTCTGATATGGCCATGATTTGTACAGAC	480
Db	421 GTGGCAGGCTTTCGCGCTTGGTACTGTTCTCTGATATGGCCATGATTTGTACAGAC	480
Qy	481 TTTTAAACCTTTATCCTACCAACAATTAAAGTATGATTGGCCCTGCATCTTAAAT	540
Db	481 TTTTAAACCTTTATCCTACCAACAATTAAAGTATGATTGGCCCTGCATCTTAAAT	540
Qy	541 GGGTGGGACGGGTCTGCGCTATGATCCTTGGAGATGACCTGCTCTCTGTTCTGTGCT	600
Db	541 GGGTGGGACGGGTCTGCGCTATGATCCTTGGAGATGACCTGCTCTCTGTTCTGTGCT	600

QY	336	GCTGTCGGGGAGACCAACAAGTGAAGGAAGGCCGATAGCCATGGGTGAGAGCAATATT	415
Db	2359	CCTGGTGGCCCCGGGAC--CCGGCCAAAGGCCCGTGTGACCTCAACGGAGAGCGTGCTCT	2455
QY	416	TCATCGTGGCAGGCTTTCGCCGCTTGATCTGTGCTCCGTATGAGCCATCAGATTGCA	475
Db	2456	ACCTGTATTGGGGGTGCTGGCGCTGAGGCCACTGTGCTGGTTCGCAACATTGTGCTCC	2515
QY	476	CAGACTTTTATAACCCCTTGTGATCCTTACCAACATTAGTAGAGTTTGGCCCTGGCATCT	535
Db	2516	GCGAGTTTATGACCCCGCTGTGCCCCGTGTGCGACAGATATACAGCTGGGCGCAGCGCTGT	2575
QY	536	TTATTGGCTGGGACAGGATCTGCCATTCATCTCGGAGAGTGCATCTGCTCTCTG	590
Db	2576	ACATTCGCTGGGCGGCCCAACCGCGCTCTCATATGTATGAGCGGCTGCTCTTTGTGCTG	2630

```

RESULT 2
US-10-765-790-158
; Sequence 158, Application US/10765790
; GENERAL INFORMATION:
; APPLICANT: Bayer Healthcare LLC
; APPLICANT: Beard, Chris
; APPLICANT: Burgess, Chris
; APPLICANT: Gannon, Allison
; APPLICANT: Harvey, Jeanne
; APPLICANT: Lechner, John F.
; APPLICANT: Li, Zheng
; TITLE OF INVENTION: Identification and Verification of Methylation Marker Sequences
; FILE REFERENCE: 1657/2035
; CURRENT APPLICATION NUMBER: US/10/765,790
; CURRENT FILING DATE: 2004-01-27
; PRIOR APPLICATION NUMBER: US 10/737,062
; PRIOR FILING DATE: 2003-12-16
; NUMBER OF SEQ ID NOS: 300
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 158
; LENGTH: 3311
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-765-790-158

```

Query Match	Best Local Match	Similarity	Score	DB	Length
26.4%	59.8%	179.8	DB 6	3311	
Matches 320	Conservative 0	Mismatches 212	Indels 3	Gaps 1	
QY	56	TGCAATTGCTGGGCTTCTCCATGAGCCCTGCTGGGCTGGGTGGGATCTGGGTGGCTTCACCG	115		
Db	2099	TGAGATCCTGGGCTGGTGGCTGTGCTGTGCTGGTGGGCTGGGGGGGTGTGATCTCGGGCTG	2158		
QY	116	CCATCCCGCAGTGGGAGATGAGACTCTTATGGGGGTGACAAATCATCTACGGCCAGGCCA	175		
Db	2159	GAGTCCCATGTGGAGAGGTGAGCCGGCTTCTGGAGCAAAATCATCTGACGGCGAGACCA	2218		
QY	176	TGTCAAGAGGAGCTGTGATGAGACTGCTTCACGACAGACAGGGAGTGAATGAGCTGCACAA	235		
Db	2219	CCTGGAAGGGGCTGTGATGTGTGTCGTGCTGTGCTGACAGACACGGGGACATGACATGACAA	2278		
QY	236	TGTACGACTGAGTGTCTGCCCTGTGTCGGGCGCTTGCAGGCCCATTCGAGCCCTTAATGTGG	295		
Db	2279	TGTACGACTGAGTGTCTGTGCTGTGCTGTGACACCGAGAGTGCAGGCGCGGCGCTTACCGTGA	2338		
QY	296	TCTCCCTGATGCTGGGCTTCCGTGGCCATGTTTGTGGCCACGATGGGACATGAAAGTGCACGC	355		
Db	2339	GCGCGTGTCTGTGCGCTTTCGTTGCGCTTTCGTGATCCCTGACGGGCGCGGACGTGCACCA	2398		
QY	356	GCTGTGGGGAGACACACAAGTGAAGAGGCCCGATATAGCCATGGGTGAGAGCATTAATT	415		
Db	2399	CCTGGTGGCCCCGGGCG---CGGCGCAAGCGCGGTGTGACCTTCACCGGAGGCGGTCT	2455		
QY	416	TCATGTGTGACAGTGTGTGCGGCTTGTGTGATGCTTCTGGTATGTGACATAGATTGTCA	475		
Db	2456	ACCTGTTTTGGGGGTGTGTGTGGGCTGTGACACTGTGTGTGTGGCCAAATGTGTGCTC	2515		

Oy 476 CAGCCTTTTAAACCCCTTTGATCCCTCAACAACTAAGTAAAGTATTGGCCCTGCACAT 535
Db 2516 GCGAGTTTTCAGACCCGCTGTGTGCGCCGAGTGCAGAAATGACAGACTGGGCGCAGGCGTGT 2575
Oy 536 TTATTGCTGTGGGACGAGTCTGCCTTAATCAATCCCTGGAGATGCACTGCTCTCCG 590
Db 2576 ACATCCGCTGGGCGGCCCAACCGCGCTGCTCAATGATAGGAGGCGTGCCTCTTTGAGCTG 2630

```

RESULT 3
US-10-765-790-1
; Sequence 1, Application US/10765790
; GENERAL INFORMATION:
; APPLICANT: Bayer Healthcare LLC
; APPLICANT: Beard, Chris
; APPLICANT: Burgess, Chris
; APPLICANT: Gannon, Allison
; APPLICANT: Harvey, Jeanne
; APPLICANT: Lechner, John F.
; APPLICANT: Li, Zheng
; TITLE OF INVENTION: Identification and Verification of Methylation Marker Sequences
; FILE REFERENCE: 1657/2035
; CURRENT APPLICATION NUMBER: US/10/765,790
; CURRENT FILING DATE: 2004-01-27
; PRIORITY APPLICATION NUMBER: US 10/737,082
; PRIORITY FILING DATE: 2003-12-16
; NUMBER OF SEQ ID NOS: 300
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 3028
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-765-790-1

```

Query Match	Best Local Similarity	22.0% ; Score 150; DB 6; Length 3028;
Matches 314; Conservative 0; Mismatches 255; Indels 1; Gaps 1	55.1% ; Pred. No. 1.2e-34;	
Qy 57	GCACTTGCTGAGGCTTCCTCCATGAGCCCTGCTGAGCTGAGGTGGGTCTGTGAGCTGCACCGC	116
Dh 1128	GAATTCCTGGGGCTGTTCTTCTTGAGTGTTGGAAT-GATGGGACAGGTGCTGTCACTGT	1186
Qy 117	CATCCCGCAGTGGAGATGAGCTCTCTTATGCGGGTGAACAATATATCAGCGCCAGGCAT	176
Dh 1187	CATGCTCAGTGGAGAGTGTGCGCTTCATTGAAAACAATCTGTGTTTGGAAAATT	1246
Qy 177	GTACAAGGGGCTGTGATGAGTCACTGTCACGACAGACGGGGATGATGAGCTGCACAAAT	236
Dh 1247	CTGGGAAGAGACTGTGATGAAATTCGTGAGGACGGCTAACTACAGATGAGTGCACAAAT	1306
Qy 237	GTACAGCTCGGTCTGCTCCCTTCGCGGGCCCTTGACAGCCACTGAGCCCTTAATGTGTGT	296
Dh 1307	CTAATGATTCCTCGTGGCTCTTTCTCCGGACCTTACAGGACGACGAGGACTGAATGTGTGC	1366
Qy 297	CTCCCTGTGCTGGGCTTCCTCGGCACATGTTGTGGCCACATAGGCATGAAATGTGACGG	356
Dh 1367	TGCTTCGTGATGTCTTCTTGGTTTCAATGATGGCCATCTTGGCATTGAATGACACAG	1426
Qy 357	CTGTGGGGAGACGACAAATGTAAGAAAGGCCCTTATAGCATGTGGTGGAGCATAAATTT	416
Dh 1427	GTGACAGGGGAGCAATAGAAAGGTGAAGGCTCAATTCTGTGACGGGTGAATCATCTTT	1486
Qy 417	CATGTGGCAGGTCTTGCCGCTTGTTGATGTGTCTCTGTGATGGCATCAGATTTGTAC	476
Dh 1487	CATCATACGGGCAATGTGTGTGTCTCACTCTGTAGCTGGGTTGCCAATGCAATCATCAG	1546
Qy 477	AGACTTTATTAACCCCTTGATCCCTACCAACATTAATATAGATTGGCCCTGCATCTT	536
Dh 1547	AGATTCTTAATCACTAATAGATGTGTGCCAAAAACGTGACGTTGGAGAGACTCTCTA	1606
Qy 537	TATTGGCTGGACAGGCTGACCTTAATCTGAGAGGTGACATCTGCTCTCTGTCTCTG	596

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 24, 2004, 05:54:20 ; Search time 3134 Seconds

(without alignments)
6488.878 Million cell updates/sec

Title: US-10-068-486a-1

Perfect score: 681

Sequence: 1 cctgaagtcctgtccaggg.....aggagatcgtgagtcgac 681

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:

1: em_estba:*
2: em_estbm:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gse_hum:*
18: em_gse_inv:*
19: em_gse_pln:*
20: em_gse_vrt:*
21: em_gse_fun:*
22: em_gse_mus:*
23: em_gse_pro:*
24: em_gse_rnd:*
25: em_gse_pmg:*
26: em_gse_vrl:*
27: em_gse_vrl:*
28: gb_gse1:*
29: gb_gse2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	634.6	93.2	866	14	CA489044 AGENCOURT
2	614.4	90.2	920	13	BU856629 AGENCOURT
3	574.4	84.3	733	12	BG747387 602704864
4	571.4	83.9	734	12	BG385787 602454019

5	555.2	81.5	936	13	BU944031	AGENCOURT
6	553	81.2	804	13 <td>BU619669 <td>UT-H-FT1-</td> </td>	BU619669 <td>UT-H-FT1-</td>	UT-H-FT1-
7	544.4	79.9	809	14 <td>CD366518 <td>UT-H-FT1-</td> </td>	CD366518 <td>UT-H-FT1-</td>	UT-H-FT1-
8	527	77.4	675	12 <td>BM824046 <td>K-EST0095</td> </td>	BM824046 <td>K-EST0095</td>	K-EST0095
9	519.6	76.3	605	12 <td>BM781639 <td>K-EST0058</td> </td>	BM781639 <td>K-EST0058</td>	K-EST0058
10	518	76.1	669	12 <td>BI753859</td> <td>603027541</td>	BI753859	603027541
11	514.4	75.5	1190	9 <td>AL547232 <td>AL547232</td> </td>	AL547232 <td>AL547232</td>	AL547232
12	513.8	75.4	543	10 <td>BF229966 <td>PM3-CS003</td> </td>	BF229966 <td>PM3-CS003</td>	PM3-CS003
13	513.8	75.4	767	12 <td>BI905712</td> <td>603166179</td>	BI905712	603166179
14	513.8	75.4	1221	11 <td>AK087296 <td>Mus muscu</td> </td>	AK087296 <td>Mus muscu</td>	Mus muscu
15	513.8	75.4	1224	11 <td>AK002924 <td>Mus muscu</td> </td>	AK002924 <td>Mus muscu</td>	Mus muscu
16	507.4	74.5	926	12 <td>BI759545</td> <td>603046958</td>	BI759545	603046958
17	506	74.3	940	13 <td>BI694715</td> <td>603347659</td>	BI694715	603347659
18	502.8	73.8	940	12 <td>BU956748 <td>AGENCOURT</td> </td>	BU956748 <td>AGENCOURT</td>	AGENCOURT
19	496.4	72.9	749	14 <td>CB851489 <td>UT-CF-EN1</td> </td>	CB851489 <td>UT-CF-EN1</td>	UT-CF-EN1
20	495	72.7	772	12 <td>BM980615 <td>UT-CF-EN1</td> </td>	BM980615 <td>UT-CF-EN1</td>	UT-CF-EN1
21	494	72.5	756	12 <td>BM975382 <td>UT-CF-EN1</td> </td>	BM975382 <td>UT-CF-EN1</td>	UT-CF-EN1
22	492.4	72.3	932	12 <td>BI408605</td> <td>602964278</td>	BI408605	602964278
23	491.8	72.0	896	12 <td>BG747546</td> <td>602704679</td>	BG747546	602704679
24	490.6	72.0	776	14 <td>CK129234 <td>AGENCOURT</td> </td>	CK129234 <td>AGENCOURT</td>	AGENCOURT
25	489.2	71.8	938	12 <td>BI106853</td> <td>602894627</td>	BI106853	602894627
26	489.2	71.8	828	12 <td>BC966566</td> <td>602833091</td>	BC966566	602833091
27	488.8	71.8	871	12 <td>BI854023</td> <td>603380915</td>	BI854023	603380915
28	484	71.1	743	14 <td>CA426782 <td>UT-H-FE1-</td> </td>	CA426782 <td>UT-H-FE1-</td>	UT-H-FE1-
29	479.6	70.4	746	9 <td>AL148864 <td>AL148864</td> </td>	AL148864 <td>AL148864</td>	AL148864
30	479	70.3	1074	9 <td>AL574569 <td>AL574569</td> </td>	AL574569 <td>AL574569</td>	AL574569
31	478.4	70.2	539	10 <td>AW952945 <td>EST365015</td> </td>	AW952945 <td>EST365015</td>	EST365015
32	477.4	70.1	904	14 <td>CA454524 <td>AGENCOURT</td> </td>	CA454524 <td>AGENCOURT</td>	AGENCOURT
33	471.8	69.3	926	13 <td>BU500215 <td>AGENCOURT</td> </td>	BU500215 <td>AGENCOURT</td>	AGENCOURT
34	471.2	69.2	892	14 <td>CA455047 <td>AGENCOURT</td> </td>	CA455047 <td>AGENCOURT</td>	AGENCOURT
35	469.8	69.0	613	12 <td>BG967711</td> <td>602833225</td>	BG967711	602833225
36	469.8	69.0	707	10 <td>BE673121</td> <td>7029702.X</td>	BE673121	7029702.X
37	466.6	68.5	562	10 <td>BE234654</td> <td>141829.MA</td>	BE234654	141829.MA
38	466.4	68.5	779	12 <td>BI654396</td> <td>603208659</td>	BI654396	603208659
39	465.4	68.3	702	12 <td>BI653083</td> <td>603300601</td>	BI653083	603300601
40	465.2	68.3	782	13 <td>BO203981 <td>UT-R-DN1-</td> </td>	BO203981 <td>UT-R-DN1-</td>	UT-R-DN1-
41	464.4	68.2	621	12 <td>BG619894</td> <td>602618081</td>	BG619894	602618081
42	464.2	68.2	878	12 <td>BC9823140</td> <td>602726481</td>	BC9823140	602726481
43	463	67.6	743	10 <td>BF531587</td> <td>602091261</td>	BF531587	602091261
44	460.4	68.0	790	10 <td>BF163497</td> <td>601771673</td>	BF163497	601771673
45	459.6	67.5	538	10 <td>BF333880</td> <td>PM3-CS003</td>	BF333880	PM3-CS003

ALIGNMENTS

RESULT 1
CA489044
LOCUS
DEFINITION AGENCOURT 10853579 MAPCL Homo sapiens cDNA clone IMAGE:6721294 5',
mRNA sequence.
ACCESSION CA489044.1 GI:24951835
VERSION
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
EST.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 866)
NTH-MGC http://mgc.ncl.nih.gov/
AUTHORS
TITLE
JOURNAL
COMMENT
NTH-MGC http://mgc.ncl.nih.gov/
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgaabs-remail.nih.gov
Tissue Procurement: Kristi A. Bglund, Ira Pastan
CDNA Library Preparation: Invitrogen Corp
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LHAM14282 row: c column: 22
High quality sequence stop: 688.

FEATURES
source

Location/Qualifiers
1. 866
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6721294"
/cell_line="ZR-75-1, MCF7, SK-BR-3, MDA-MB-231,
hTERT-HME1, LNCaP"
/lab_host="EMD105"
/clone_id="MAPcl"
/note="Vector: pCMV-Sport6; Site_1: EcoRV; Site_2: Not I;
Subtracted with brain, liver, lung, kidney and muscle.
Directionally cloned. Priming method: oligo-dT. Average
insert size: 1800 bp. Library amplification: 26,000 fold.
Kristi A. Eglund, James J. Vincent, Robert Strassberg,
Bundkook Lee & Ira Pastan. Discovery of new breast
cancer genes encoding membrane and secreted proteins.
Manuscript submitted."

ORIGIN

Query Match 93.2%; Score 634.6; DB 14; Length 866;
Best Local Similarity 99.4%; Pred. No. 5.1e-125;
Matches 637; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

QY 35 CGCGCATGCGCAATGCGGCGCTGCGAGTGTGCGGCTTCCATGAGCCCTGCTGGGCTGG 94
DB 113 CGGAATGCGCAATGCGGCGCTGCGAGTGTGCGGCTTCCATGAGCCCTGCTGGGCTGG 172
QY 95 TGGGCTGCTGCGCTGCGACCGCGCATCCCGAGTGGCAGATGAGCTCCTATGCGGGTACA 154
DB 173 TGGGCTGCTGCGCTGCGACCGCGCATCCCGAGTGGCAGATGAGCTCCTATGCGGGTACA 232
QY 155 ACATATACAGCGCCCAAGCGCATGTACAGAGGCGCTGTGATGAGTGTGCTGACGAGACA 214
DB 233 ACATATACAGCGCCCAAGCGCATGTACAGAGGCGCTGTGATGAGTGTGCTGACGAGACA 292
QY 215 CGGGGATGATGAGCTGCAAAATGTACAGTGTGCTGCGCTGCGCGGCTTGGAGG 274
DB 293 CGGGGATGATGAGCTGCAAAATGTACAGTGTGCTGCGCTGCGCGGCTTGGAGG 352
QY 275 CCATCGAGCCCTTAATGCTGTCTGCTGCTGCTGCGGCTTCTGCGCATGTTTGGCCA 334
DB 353 CCATCGAGCCCTTAATGCTGTCTGCTGCTGCTGCGGCTTCTGCGCATGTTTGGCCA 412
QY 335 CGATGCGCATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 394
DB 413 CGATGCGCATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 472
QY 395 CCATGCGCATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 454
DB 473 CCATGCGCATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 532
QY 455 GGTATGCGCATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 514
DB 533 GGTATGCGCATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 592
QY 515 ATGAGTTTGGCCCTTCAATTTTATGCTGTGCGGAGGCTGCTGCTGCTGCTGCTGCTG 574
DB 593 ATGAGTTTGGCCCTTCAATTTTATGCTGTGCGGAGGCTGCTGCTGCTGCTGCTGCTG 652
QY 575 GTGCACTGCTCTCCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 634
DB 653 GTGCACTGCTCTCCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 712
QY 635 GCTCTTACCTTAAGTCAAACTCTTCAAGAGATGTTGA 675
DB 713 GCTCTTACCTTAAGTCAAACTCTTCAAGAGATGTTGA 753

```

RESULT 2
BUS856629 920 bp mRNA linear EST 16-OCT-2002
LOCUS BUS856629 920 bp mRNA linear EST 16-OCT-2002
DEFINITION AGNCOURT_10474057 NIH_MGC_107 Homo sapiens cDNA clone

FEATURES
source

IMAGE:6646374 5', mRNA sequence.
BUS856629
EST. GI:24041619
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE
1 (bases 1 to 920)
NIH-MGC http://mhc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
CONTACT: Robert Strassberg, Ph.D.
Email: cga@hs-ni.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Cloned through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LNC2885 row: b column: 06
High quality sequence stop: 676.

FEATURES
source

1. 920
Location/Qualifiers

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6646374"
/tissue="adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_id="NIH MGC 107"
/note="Organ: breast; Vector: pOT7; Site_1: EcoRI;
Site_2: XhoI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGGCAGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

ORIGIN

Query Match 90.2%; Score 614.4; DB 13; Length 920;
Best Local Similarity 99.8%; Pred. No. 1.1e-120;
Matches 615; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 60 GTTGTGCGCTTCTCAATGCGGCGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCG 119
DB 1 GTTGTGCGCTTCTCAATGCGGCGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCG 60
QY 120 CCGCAGTGGCAGATGAGTCTCTATGCGGCTGACACATCATCATCAGGCGCCATGTA 179
DB 61 CCGCAGTGGCAGATGAGTCTCTATGCGGCTGACACATCATCATCAGGCGCCATGTA 120
QY 180 CAAGGCGCTGTGATGAGTGTGCTGACGACAGACAGCGGAGATGAGTGTGAAATGTA 239
DB 121 CAAGGCGCTGTGATGAGTGTGCTGACGACAGACAGCGGAGATGAGTGTGAAATGTA 180
QY 240 CCACTGCTGCTGCGCTGCTCCGCGGCTTGAAGGCACTCGAGCCCTAATGCTGCTC 299
DB 181 CCACTGCTGCTGCGCTGCTCCGCGGCTTGAAGGCACTCGAGCCCTAATGCTGCTC 240
QY 300 CTTGTGCTGCGCTTCTGCGCATGTTTGTGCGCATGAGGATGAGTGTGAGCGGCTG 359
DB 241 CTTGTGCTGCGCTTCTGCGCATGTTTGTGCGCATGAGGATGAGTGTGAGCGGCTG 300
QY 360 TGGGAGAGCAGCAAAAGTGAAGAGAGCGGCTATAGCATGAGTGTGAGGATATTTTCA 419
DB 301 TGGGAGAGCAGCAAAAGTGAAGAGAGCGGCTATAGCATGAGTGTGAGGATATTTTCA 360
QY 420 CTTGTGAGGCTTTGGCGCTTGTGTAGCTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTG 479
DB 361 CTTGTGAGGCTTTGGCGCTTGTGTAGCTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTG 420

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 25, 2004, 02:34:14 ; Search time 492 Seconds

(without alignments)
90.556 Million cell1 updates/sec

Title: US-10-068-486a-2

Sequence: 1 MANWGLQLLGFSMALIGWV.....KAGRAPRSYKSNSSKEYV 211

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 809742 seqs, 21153259 residues

Total number of hits satisfying chosen parameters: 809742

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA.*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTOS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1109	100.0	211	14	US-10-068-486a-2
2	1095	98.7	211	14	US-09-471-749-7
3	1095	98.7	211	14	US-10-176-847-84
4	1095	98.7	211	15	US-10-236-031B-14
5	1095	98.7	247	14	US-10-106-698-6273
6	1091	98.4	210	9	US-09-938-390A-36
7	1091	98.4	211	15	US-10-341-434-77
8	1091	98.4	247	15	US-10-264-049-3053
9	812	73.2	194	9	US-09-925-301-999
10	705	63.6	211	9	US-09-886-683A-4
11	705	63.6	211	9	US-09-978-295A-270
12	705	63.6	211	9	US-09-978-697-270
13	705	63.6	211	9	US-09-978-192A-270
14	705	63.6	211	9	US-09-999-832A-270
15	705	63.6	211	10	US-09-978-189-270

16	705	63.6	211	10	US-09-978-608A-270	Sequence 270, App
17	705	63.6	211	10	US-09-978-585A-270	Sequence 270, App
18	705	63.6	211	10	US-09-978-191A-270	Sequence 270, App
19	705	63.6	211	10	US-09-978-403A-270	Sequence 270, App
20	705	63.6	211	10	US-09-978-564A-270	Sequence 270, App
21	705	63.6	211	10	US-09-999-833A-270	Sequence 270, App
22	705	63.6	211	10	US-09-981-915A-270	Sequence 270, App
23	705	63.6	211	10	US-09-978-824-270	Sequence 270, App
24	705	63.6	211	10	US-09-918-585A-270	Sequence 270, App
25	705	63.6	211	10	US-09-978-423A-270	Sequence 270, App
26	705	63.6	211	10	US-09-978-188A-270	Sequence 270, App
27	705	63.6	211	10	US-09-899-830A-270	Sequence 270, App
28	705	63.6	211	10	US-09-978-757A-270	Sequence 270, App
29	705	63.6	211	10	US-09-978-187B-270	Sequence 270, App
30	705	63.6	211	10	US-09-978-643A-270	Sequence 270, App
31	705	63.6	211	10	US-09-978-375A-270	Sequence 270, App
32	705	63.6	211	10	US-09-978-298A-270	Sequence 270, App
33	705	63.6	211	10	US-09-978-188A-270	Sequence 270, App
34	705	63.6	211	10	US-09-978-681A-270	Sequence 270, App
35	705	63.6	211	10	US-09-978-194A-270	Sequence 270, App
36	705	63.6	211	10	US-09-999-829A-270	Sequence 270, App
37	705	63.6	211	10	US-09-978-544A-270	Sequence 270, App
38	705	63.6	211	10	US-09-978-829A-270	Sequence 270, App
39	705	63.6	211	10	US-09-978-544A-270	Sequence 270, App
40	705	63.6	211	10	US-09-978-665A-270	Sequence 270, App
41	705	63.6	211	13	US-10-052-586-102	Sequence 270, App
42	705	63.6	211	13	US-10-105-929-4	Sequence 102, App
43	705	63.6	211	13	US-10-174-590-102	Sequence 102, App
44	705	63.6	211	14	US-10-176-758-102	Sequence 102, App
45	705	63.6	211	14	US-10-175-737-102	Sequence 102, App

ALIGNMENTS

RESULT 1
US-10-068-486a-2
Sequence 2, Application US/10068486a
Publication No. US20030148939A1
GENERAL INFORMATION:
APPLICANT: Genzyme Corporation
TITLE OF INVENTION: COMPOSITIONS AND METHODS RELATED TO CLAUDIN-7
FILE REFERENCE: G02217C
CURRENT FILING DATE: 2002-02-05
PRIOR APPLICATION NUMBER: US/10/068,486A
PRIOR FILING DATE: 2000-08-07
PRIOR APPLICATION NUMBER: PCT/US00/24174
PRIOR FILING DATE: 1999-08-06
NUMBER OF SEQ ID NOS: 2
SOFTWARE: PatentIn version 3.1
SEQ ID NO 2
LENGTH: 211
TYPE: PRT
ORGANISM: Homo sapiens
US-10-068-486a-2

Query Match 100.0%; Score 1109; DB 14; Length 211;
Best Local Similarity 100.0%; Pred No. 6; 7e-111;
Matches 211; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MANWGLQLLGFSMALIGWGLVACTAIPOMQSSYAGDNIITQAMYKSLIMDCVQSTG 60
DB 1 MANWGLQLLGFSMALIGWGLVACTAIPOMQSSYAGDNIITQAMYKSLIMDCVQSTG 60
QY 61 MMSCKRYDSTLALSAALQTRALMVSVLGLFAMVAATMGKCTCGGDDKXKARIM 120
DB 61 MMSCKRYDSTLALSAALQTRALMVSVLGLFAMVAATMGKCTCGGDDKXKARIM 120
QY 121 GGGIIFVAGLALVACSWYSGQIVTDFFNPLIPITIKYKFGPAIFIGWAGSLVILGGA 180
DB 121 GGGIIFVAGLALVACSWYSGQIVTDFFNPLIPITIKYKFGPAIFIGWAGSLVILGGA 180

QY 181 L1SCCPGNEBKAGYRAPRSYPKSNSKEYV 211
DB 181 L1SCCPGNEBKAGYRAPRSYPKSNSKEYV 211

RESULT 2

US-09-471-749-7
Sequence 7, Application US/09471749
Publication No. US20030124113A1
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Corley, Neil C.
APPLICANT: Guegler, Karl J.
APPLICANT: Paterson, Chandra
APPLICANT: Baughn, Mariah
TITLE OF INVENTION: HUMAN APOPTOSIS ASSOCIATED PROTEINS
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Dr.
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/471,749
FILING DATE:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 09/078,402
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Geirone, Michael C.
REGISTRATION NUMBER: 39,132
REFERENCE/DOCKET NUMBER: PF-0519 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 211 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: OVARNOT03
CLONE: 2059022
US-09-471-749-7

Query Match 98.7%; Score 1095; DB 10; Length 211;
Best Local Similarity 99.5%; Pred. No. 2.1e-109;
Matches 210; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MANMGLQILGFSSMALLGWVGLVACTAIPQWOMSSYAGDNITTAQAMTKGLMMDCVTOSTG 60
DB 1 MANSGLQILGFSSMALLGWVGLVACTAIPQWOMSSYAGDNITTAQAMTKGLMMDCVTOSTG 60
QY 61 MMSCKMYDSVLAALQATRALMVSIVLGFAMFVATMGKCTRCGGDDKVKKARIAM 120
DB 61 MMSCKMYDSVLAALQATRALMVSIVLGFAMFVATMGKCTRCGGDDKVKKARIAM 120
QY 121 GGGIIFIVAGLAALVACSWYGHQIVTDYFNPLIPNIXEFGPAIFIGWAGSALVILGGA 180
DB 121 GGGIIFIVAGLAALVACSWYGHQIVTDYFNPLIPNIXEFGPAIFIGWAGSALVILGGA 180
QY 181 L1SCCPGNEBKAGYRAPRSYPKSNSKEYV 211
DB 181 L1SCCPGNEBKAGYRAPRSYPKSNSKEYV 211

RESULT 3

US-10-176-847-84
Sequence 84, Application US/10176847
Publication No. US20030068636A1
GENERAL INFORMATION:
APPLICANT: Velby, Peter Ole
TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR
IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF BREAST
TITLE OF INVENTION: AND OVARIAN CANCER
FILE REFERENCE: MRI-039
CURRENT APPLICATION NUMBER: US/10/176,847
CURRENT FILING DATE: 2002-06-21
NUMBER OF SEQ ID NOS: 112
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 84
LENGTH: 211
TYPE: PRT
ORGANISM: Homo sapiens
US-10-176-847-84

Query Match 98.7%; Score 1095; DB 14; Length 211;
Best Local Similarity 99.5%; Pred. No. 2.1e-109;
Matches 210; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MANMGLQILGFSSMALLGWVGLVACTAIPQWOMSSYAGDNITTAQAMTKGLMMDCVTOSTG 60
DB 1 MANSGLQILGFSSMALLGWVGLVACTAIPQWOMSSYAGDNITTAQAMTKGLMMDCVTOSTG 60
QY 61 MMSCKMYDSVLAALQATRALMVSIVLGFAMFVATMGKCTRCGGDDKVKKARIAM 120
DB 61 MMSCKMYDSVLAALQATRALMVSIVLGFAMFVATMGKCTRCGGDDKVKKARIAM 120
QY 121 GGGIIFIVAGLAALVACSWYGHQIVTDYFNPLIPNIXEFGPAIFIGWAGSALVILGGA 180
DB 121 GGGIIFIVAGLAALVACSWYGHQIVTDYFNPLIPNIXEFGPAIFIGWAGSALVILGGA 180
QY 181 L1SCCPGNEBKAGYRAPRSYPKSNSKEYV 211
DB 181 L1SCCPGNEBKAGYRAPRSYPKSNSKEYV 211

RESULT 4

US-10-236-031B-14
Sequence 14, Application US/10236031B
Publication No. US20030219760A1
GENERAL INFORMATION:
APPLICANT: Gordon, Gavin J.
APPLICANT: Jensen, Roderick V.
APPLICANT: Gullane, Steven R.
APPLICANT: Bueno, Raphael
TITLE OF INVENTION: Diagnostic and Prognostic Tests
FILE REFERENCE: B00801/70265 (JRV/JAV)
CURRENT APPLICATION NUMBER: US/10/236,031B
CURRENT FILING DATE: 2002-09-05
PRIORITY APPLICATION NUMBER: US 60/317,389
PRIORITY FILING DATE: 2001-09-05
PRIORITY APPLICATION NUMBER: US 60/407,431
PRIORITY FILING DATE: 2002-08-30
NUMBER OF SEQ ID NOS: 102
SOFTWARE: PatentIn version 3.1
SEQ ID NO 14
LENGTH: 211
TYPE: PRT
ORGANISM: Homo sapiens
US-10-236-031B-14

Query Match 98.7%; Score 1095; DB 15; Length 211;
Best Local Similarity 99.5%; Pred. No. 2.1e-109;
Matches 210; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MANMGLQILGFSSMALLGWVGLVACTAIPQWOMSSYAGDNITTAQAMTKGLMMDCVTOSTG 60

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 25, 2004, 02:18:44 / Search time 72 Seconds

(without alignments)
281.895 Million cell updates/sec

Title: US-10-068-486a-2

Sequence: 1109
1 MANWGLQLGFSMALLGWVG.....KAGYRAPRSYKSNSEKEYV 211

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database:

1: PIR 78: *
2: PIR: *
3: PIR: *
4: PIR: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	516	46.5	280	A39484	androgen-withdrawn
2	114.5	10.3	160	UN0503	peripheral myelin
3	114.5	10.3	160	S21721	growth arrest-spec
4	112.5	10.1	160	A41144	growth arrest-rela
5	101	9.1	167	T16454	hypothetical prote
6	98	8.8	425	A97688	hypothetical 45.5K
7	98	8.8	425	AF2913	conserved hypothet
8	97.5	8.8	959	T25704	hypothetical prote
9	95	8.6	385	E83506	probable MFS trans
10	94	8.5	173	A49182	integral membrane
11	93	8.4	402	H81296	probable integral
12	92.5	8.3	144	A36324	growth arrest-spec
13	92.5	8.3	157	G02355	tumor-associated m
14	92	8.3	411	A87390	transporter, proba
15	90.5	8.2	455	1 WQBST	phosphotransferase
16	90	8.1	442	C82985	ammonium transport
17	89.5	8.1	472	F82639	resistance protein
18	89	8.0	409	B85735	probable membrane
19	89	8.0	417	C81084	probable integral
20	89	8.0	421	E90883	hypothetical prote
21	89	8.0	428	D64895	probable membrane
22	89	8.0	436	C69903	proline permease h
23	87.5	7.9	428	B64081	fructose permease h
24	87	7.8	167	JC5044	epithelial membran
25	87	7.8	307	D36868	copp homolog - xan
26	87	7.8	338	AE3334	metal chelate tran
27	86.5	7.8	342	E71502	probable apolipop
28	86.5	7.8	744	T13429	hypothetical prote
29	86	7.8	173	A48300	lens membrane prot

30	86	7.8	977	2	S53302	H+-exporting ATPas
31	85.5	7.7	242	2	T15361	hypothetical prote
32	85	7.7	477	2	B28988	hypothetical prote
33	84	7.6	401	2	B82951	probable MFS trans
34	84	7.6	426	2	B98106	hypothetical prote
35	83.5	7.5	445	2	B83885	cytosine permease
36	83	7.5	279	2	T29455	hypothetical prote
37	83	7.5	425	2	JC4179	vasopressin recept
38	83	7.5	441	2	G82253	conserved hypothet
39	82.5	7.4	246	2	D97074	terc family protei
40	82.5	7.4	407	2	AC2721	MFS permease (limp
41	82.5	7.4	407	2	G97502	probable MFS trans
42	82	7.4	352	2	A97541	hypothetical prote
43	82	7.4	352	2	AC2760	sodium bile acid s
44	82	7.4	411	2	G95241	MATE efflux family
45	82	7.4	437	2	H81041	citrate transporte

ALIGNMENTS

RESULT 1

A39484
androgen-withdrawal apoptosis protein RVPI, prostatic - rat
C/Species: Rattus norvegicus (Norway rat)
C/Date: 03-Aug-1992 #sequence_revision 03-Aug-1992 #text_change 04-Mar-2000
C/Accession: A39484
R/Briehl, M.M./Miesfeld, R.L.
Mol. Endocrinol. 5, 1381-1388, 1991
A>Title: Isolation and characterization of transcripts induced by androgen withdrawal at
A/Reference number: A39484; MUID:92130987; PMID:1723140
A/Accession: A39484
A/Molecule type: mRNA
A/Residues: 1-280 <BRI>
A/Cross-references: GB:M74067; NID:G205857; PIDN:AAA41760.1; PID:G205858
A/Genetics:
A/Gene: RVP.1
C/Supfamily: rat androgen-withdrawal apoptosis protein RVPI

Query Match 46.5%; Score 516; DB 2; Length 280;
Best Local Similarity 49.2%; Pred. NO. 4.8e-41;
Matches 98; Conservative 39; Mismatches 56; Indels 6; Gaps 3;

QY	6	IQLLGFSWALIGWGLVACTAIPOWSSSYAGNIITIAQAWYKGLMDCVTSQGMNSCK	65
DB	5	LEIGTSLAVGWICTIVCCALPMKRVSAFSGSITITQITWEGLMNVCV-QSTGCMQCK	63
QY	66	MYDSVLAISALQATRALMVSVLVGLAMFVATMGKCTRCGSDDYKARIMGGGII	125
DB	64	MYDSLALPQDLQARALIVSIIIAAFGLVALVGAQCTNC-VQDETAKAKITIVAGVL	122
QY	126	FIVAGIALVACSVYGHQIVDFYNPILPTNIKXFEFGAPFPGWAGSLVIGALISCS	185
DB	123	FLIAVLVTVSVSSANTITIDFNPVPEAKREMGTLVGNAAAALLQLGALLCCS	182
QY	186	CPGNE---SRAGYRAPRS	200
DB	183	CPPREKAPTKILYSAPRS	201

RESULT 2

UN0503
peripheral myelin protein 22 - human
N/Alternate names: Charcot-Marie-Tooth; GNS-3 protein; growth arrest-specific protein 3;
C/Species: Homo sapiens (man)
C/Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 20-Jun-2000
C/Accession: UN0503; JCI190; A56697; S25537
R/Idoni, P.; Martinotti, A.; Colombo, M.P.; Schneider, C.
Gene 126, 289-290, 1993
A>Title: Sequence of human GNS3/PM22 full-length cDNA.
A/Reference number: UN0503; MUID:93246261; PMID:8482547
A/Accession: UN0503
A/Molecule type: mRNA

A/Residues: 1-160 <EDO>
 A/Cross-references: GB:L03203; NID:9182984; PIDN:AA58495.1; PID:9182985
 R.Hayasaka, K.; Himuro, M.; Nango, K.; Sato, M.; Mura, M.; Uemura, K.; Takahashi, E.;
 Biochem. Biophys. Res. Commun. 186, 827-831, 1992
 A/Title: Isolation and sequence determination of cDNA encoding PMP-22 (PAC-II/SR13/GAS-3)
 A/Reference number: J01190; MUID:92360032; PMID:1497668
 A/Accession: J01190
 A/Molecule type: mRNA
 A/Residues: 1-160 <HAY>
 A/Cross-references: GB:D11428; NID:9220009; PIDN:BA01995.1; PID:9220010
 A/Experimental source: fetus spinal cord
 R.Valentin, J., J., Baas, F., Wolterman, R.A.; Hoogenbijk, J.E.; van den Bosch, N.H.A.; Z
 Nature Genet. 2, 288-291, 1992
 A/Title: Identical point mutations of PMP-22 in Trembler-J mouse and Charcot-Marie-Tooth
 A/Reference number: A56697; MUID:92365161; PMID:1303281
 A/Accession: A56697
 A/Status: preliminary
 A/Molecule type: mRNA
 A/Residues: 1-160 <VAL>
 A/Cross-references: GB:S61788
 R.Colombo, M.
 submitted to the EMBL Data Library, May 1992
 A/Reference number: S25537
 A/Accession: S25537
 A/Status: preliminary
 A/Molecule type: mRNA
 A/Residues: 61-160 <COL>
 A/Cross-references: EMBL:X65968; NID:931652; PIDN:CAA6781.1; PID:931653
 C/Genetic: 8
 A/Genes: GDB:PMP22; GAS3
 A/Cross-references: GDB:134190
 A/Map position: 17p12-17p11.2
 C/Superfamily: growth arrest-specific protein
 C/Keywords: myelin; transmembrane protein

Query Match 10.3%; Score 114.5; DB 2; Length 160;
 Best Local Similarity 23.9%; Pred. No. 0.0016;
 Matches 42; Conservative 28; Mismatches 77; Indels 29; Gaps 5;

QY 6 LQLGFSMALGCVGLVACTAIPQWMSYAGDNIITQAWYKGLMDCVTQSTGMSCK 65
 DB 6 LSTIYLHVAVL--VLFSTIVSQW-----LVGNHRTDLMQCTTSALGAVQ-H 52
 QY 66 MYDSVIALSALQATRALMVSLVIGFLMFVATMGKCTRCGGDDKXKARIMGGGII 125
 DB 54 FSSSP---NEWLQSVQATMILSVISLFLFCQFLTKRGFRYIT-----GVF 102
 QY 126 FIVAGLALVACSWGHQIVTDFYNPILPTNIKXEPGPAIFIGWAGSALVIL 181
 DB 103 QILAGLCWMSAAIYTVHSEWH-----VNNYYSYGFATILMVAFPLALL 148

RESULT 3
 S21721
 growth arrest-specific protein - mouse
 C/Species: Mus musculus (house mouse)
 C/Date: 22-Nov-1993 #sequence_revision 10-Nov-1993 #text_change 13-Sep-1998
 A/Accession: S21721
 R.Suter, U.; Welcher, A.A.; Oezcelik, T.; Snipes, G.J.; Kosaras, B.; Franke, U.; Billi
 Nature 356, 241-244, 1992
 A/Title: Trembler mouse carries a point mutation in a myelin gene.
 A/Reference number: S21721; MUID:92204231; PMID:1552943
 A/Accession: S21721
 A/Status: preliminary
 A/Molecule type: mRNA
 A/Residues: 1-160 <SUT>
 C/Superfamily: growth arrest-specific protein
 C/Keywords: transmembrane protein

Query Match 10.3%; Score 114.5; DB 2; Length 160;
 Best Local Similarity 22.8%; Pred. No. 0.0016;
 Matches 42; Conservative 28; Mismatches 69; Indels 45; Gaps 7;

QY 6 LQLGFSMALGCVGLVACTAIPQWMSYAGDNIITQAWYKGLMDCVTQSTGMSCK 65
 DB 6 LGILFHLHIAVL--VLFSTIVSQW-----LVGNHRTDLMQCTTSALGAVQ-H 52
 QY 66 MYDSVIALSALQATRALMVSLVIGFLMFVATMGKCTRCGGDDKXKARIMGGGII 125
 DB 53 CYSS--SVSEWLOSVQATMILSVISLFLFCQFLTKRGFRYIT-----GVF 102
 QY 126 FIVAGLALVACSWGHQIVTDFYNPILPTNIKXEPGPAIFIGWAGSALVIL 177
 DB 103 QILAGLCWMSAAIYTVHSEWH-----VNNYYSYGFATILMVAFPLALL 148

Query Match 10.1%; Score 112.5; DB 2; Length 160;
 Best Local Similarity 22.3%; Pred. No. 0.0024;
 Matches 41; Conservative 30; Mismatches 68; Indels 45; Gaps 7;

QY 6 LQLGFSMALGCVGLVACTAIPQWMSYAGDNIITQAWYKGLMDCVTQSTGMSCK 65
 DB 6 LGILFHLHIAVL--VLFSTIVSQW-----LVGNHRTDLMQCTTSALGAVQ-H 52
 QY 66 MYDSVIALSALQATRALMVSLVIGFLMFVATMGKCTRCGGDDKXKARIMGGGII 125
 DB 53 CYSS--SVSEWLOSVQATMILSVISLFLFCQFLTKRGFRYIT-----GVF 102
 QY 126 FIVAGLALVACSWGHQIVTDFYNPILPTNIKXEPGPAIFIGWAGSALVIL 177
 DB 103 QILAGLCWMSAAIYTVHSEWH-----VNNYYSYGFATILMVAFPLALL 148

RESULT 4
 A41144
 growth arrest-related myelin protein precursor, sciatic nerve - rat
 N/Alternate names: Schwann cell membrane glycoprotein, SAG
 C/Species: Rattus norvegicus (Norway rat)
 C/Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 29-Sep-1999
 A/Accession: A41144; S18550; A44826
 R.Welcher, A.A.; Suter, U.; De Leon, M.; Snipes, G.J.; Shooter, E.M.
 Proc. Natl. Acad. Sci. U.S.A. 88, 7195-7199, 1991
 A/Title: A myelin protein is encoded by the homologue of a growth arrest-specific gene.
 A/Reference number: A41144; MUID:91334432; PMID:1714591
 A/Accession: A41144
 A/Molecule type: mRNA
 A/Residues: 1-160 <WEL>
 A/Cross-references: GB:M69139; NID:9207063; PIDN:AAA73063.1; PID:9207064
 R.Spreyer, P.; Kuhn, G.; Hanemann, C.O.; Gillen, C.; Schaal, H.; Kuhn, R.; Lemke, G.; M
 EMBO J. 10, 3661-3668, 1991
 A/Title: Axon-regulated expression of a Schwann cell transcript that is homologous to a
 A/Reference number: S18550; MUID:92037527; PMID:1935894
 A/Accession: S18550
 A/Status: preliminary
 A/Molecule type: mRNA
 A/Residues: 1-29, 'R', 31-160 <SPR>
 A/Cross-references: EMBL:X62431; NID:955903; PIDN:CAA4297.1; PID:955904
 R.Diepert, W.E.; O'Neill, A.; Magnoni, G.; Wollmann, R.L.; Heintz, R.L.; Zucher-N
 J. Neurosci. 12, 2177-2185, 1992
 A/Title: SAG: a Schwann cell membrane glycoprotein.
 A/Reference number: A44826; MUID:92300442; PMID:1376775
 A/Accession: A44826
 A/Status: preliminary
 A/Molecule type: protein
 A/Residues: 1-21, 'K', 23-25, 'X', 27, 'R', 29-30 <DIB>
 A/Note: sequence extracted from NCBI backbone (NCBI:106581)
 C/Comment: The predicted signal sequence may not be cleaved in some cell types.
 C/Superfamily: growth arrest-specific protein
 C/Keywords: glycoprotein; Schwann cell; transmembrane protein

Query Match 10.1%; Score 112.5; DB 2; Length 160;
 Best Local Similarity 22.3%; Pred. No. 0.0024;
 Matches 41; Conservative 30; Mismatches 68; Indels 45; Gaps 7;

QY 6 LQLGFSMALGCVGLVACTAIPQWMSYAGDNIITQAWYKGLMDCVTQSTGMSCK 65
 DB 6 LGILFHLHIAVL--VLFSTIVSQW-----LVGNHRTDLMQCTTSALGAVQ-H 52
 QY 66 MYDSVIALSALQATRALMVSLVIGFLMFVATMGKCTRCGGDDKXKARIMGGGII 125
 DB 53 CYSS--SVSEWLOSVQATMILSVISLFLFCQFLTKRGFRYIT-----GVF 102
 QY 126 FIVAGLALVACSWGHQIVTDFYNPILPTNIKXEPGPAIFIGWAGSALVIL 177
 DB 103 QILAGLCWMSAAIYTVHSEWH-----VNNYYSYGFATILMVAFPLALL 148

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 25, 2004, 01:10:04 ; Search time 25 Seconds

(without alignments)
439,472 Million cell updates/sec

Title: US-10-068-486a-2

Sequence: 1 MAMMGLQLGFSMALLGWC.....KAGTRAPRSYKSNSEKEY 211

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : WiseProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	1095	98.7	211 1	CLD7_HUMAN
2	1023	92.2	211 1	CLD7_MOUSE
3	934	84.2	211 1	CLD7_RAT
4	705	63.6	211 1	CLD1_HUMAN
5	695	62.7	211 1	CLD1_RAT
6	694	62.6	211 1	CLD1_MOUSE
7	670	60.4	215 1	CLD5_MOUSE
8	631	56.9	211 1	CLD5_MOUSE
9	628	56.6	211 1	CLD5_MOUSE
10	572.5	51.6	210 1	CLD4_MOUSE
11	541	48.8	209 1	CLD4_MOUSE
12	540	48.7	220 1	CLD3_HUMAN
13	539	48.6	218 1	CLD3_HUMAN
14	538	48.5	209 1	CLD4_MOUSE
15	533.5	48.1	219 1	CLD3_MOUSE
16	533.5	48.1	219 1	CLD3_MOUSE
17	533	48.1	219 1	CLD3_MOUSE
18	523.5	47.2	208 1	CLD3_MOUSE
19	522	47.1	217 1	CLD3_MOUSE
20	515.5	46.5	220 1	CLD6_MOUSE
21	513.5	46.3	219 1	CLD6_MOUSE
22	508	45.8	218 1	CLD5_MOUSE
23	507.5	45.8	218 1	CLD5_MOUSE
24	504.5	45.5	218 1	CLD5_MOUSE
25	484	43.6	206 1	CLD5_MOUSE
26	474	42.7	225 1	CLD8_MOUSE
27	470	42.4	239 1	CLD8_MOUSE
28	469	42.3	225 1	CLD8_MOUSE
29	468.5	42.2	239 1	CLD8_MOUSE
30	444	40.0	224 1	CLD8_MOUSE
31	431.5	38.9	230 1	CLD2_HUMAN
32	427.5	38.5	230 1	CLD2_HUMAN
33	421.5	38.0	230 1	CLD2_MOUSE

34	410	37.0	228 1	CLD4_HUMAN
35	408	36.8	219 1	CLD4_HUMAN
36	402	36.2	228 1	CLD4_HUMAN
37	401.5	36.2	231 1	CLD4_MOUSE
38	385.5	34.8	227 1	CLD4_MOUSE
39	373.5	33.7	264 1	CLD1_MOUSE
40	365	32.9	261 1	CLD1_MOUSE
41	298	26.9	203 1	CLD1_HUMAN
42	280.5	25.3	220 1	CLD1_HUMAN
43	280	25.2	207 1	CLD1_MOUSE
44	277.5	25.0	207 1	CLD1_MOUSE
45	273	24.6	207 1	CLD1_MOUSE

ALIGNMENTS

RESULT 1
ID CLD7_HUMAN STANDARD; PRT; 211 AA.
AC 095471; 095VNO;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE claudin-7.
GN CLD7.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Colon adenocarcinoma;
RL Keen T.J.;
RT Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=22386257; PubMed=12477932.
RA Strassberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shemmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diachenko L., Marstina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ueda T.B., Toshiyuki S., Carrinci P., Prange C.,
RA Rana S.S., Loguercio N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwen P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Halik S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kertman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dixon M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Buterfield V.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC - FUNCTION: Component of tight junction (TJ) strands.
CC - SUBCELLULAR LOCATION: Integral membrane protein.
CC - SIMILARITY: Belongs to the claudin family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by, and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC EMBL; AJ011497; CA09626.1; -
CC EMBL; BC001055; AA01055.1; -

DR Genew; HGNC:2049; CLDN7.
 DR GO; GO:0016021; C: integral to membrane; TAS.
 DR GO; GO:0005923; C: tight junction; TAS.
 DR InterPro; IPR006187; Claudin.
 DR InterPro; IPR006188; Claudin_reg.
 DR InterPro; IPR004031; PMP22_Claudin.
 DR Pfam; PF00822; PMP22_Claudin; 1.
 DR PRINTS; PR01346; CLAUDIN.
 DR PROSITE; PS01346; CLAUDIN; 1.
 DR TIGHT_JUNCTION; Transmembrane; Polymorphism.
 FT TRANSMEM 8 28 POTENTIAL.
 FT TRANSMEM 82 102 POTENTIAL.
 FT TRANSMEM 118 138 POTENTIAL.
 FT TRANSMEM 161 181 POTENTIAL.
 FT VARIANT 197 197 A->V (1n dbSNP:4562).
 FT CONFLICT 133 133 A->T (1n REF. 2).
 FT SEQUENCE 211 AA; 22390 MW; 7F3C01B963D912E1 CRC64;

Query Match 98.7%; Score 1095; DB 1; Length 211;
 Best Local Similarity 99.5%; Pred. No. 1,3e-87;
 Matches 210; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MANMGQLIGFSGMALGIVGACVTAIPQMOMSSYAGDNIITAOQMYKGLMDCVTOSTG 60
 DB 1 MANSGLQLIGFSGMALGIVGACVTAIPQMOMSSYAGDNIITAOQMYKGLMDCVTOSTG 60
 QY 61 MNSCKMYDSVLAISALQATRALMNVSVLGLAFVATMGKCTRCGGDDVYKARIAM 120
 DB 61 MNSCKMYDSVLAISALQATRALMNVSVLGLAFVATMGKCTRCGGDDVYKARIAM 120
 QY 121 GGGIIFIVAGLAALVACSWYGHQIVTDFYNPILPTNIKEFGPAIFIGAGSALVILGGA 180
 DB 121 GGGIIFIVAGLAALVACSWYGHQIVTDFYNPILPTNIKEFGPAIFIGAGSALVILGGA 180
 QY 181 LLSGSCPGNESKAGRAPRSYKSNSSKEYV 211
 DB 181 LLSGSCPGNESKAGRAPRSYKSNSSKEYV 211

RESULT 2
 CLD7_MOUSE STANDARD; PRT; 211 AA.
 AC Q92261;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Claudin-7.
 GN CLDN7.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99110921; PubMed=9892664;
 RT "Claudin multigene family encoding four-transmembrane domain protein
 components of tight junction strands."
 RL Proc. Natl. Acad. Sci. U.S.A. 96:511-516(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX STRAIN=FVB/N; Tissue=Colon; and Mammary gland;
 RC MEDLINE=22389257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner K.H., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buelow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Tishiyuki S., Carrion R.P., Prange C.C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abraham R.D., Muliyil S.J.,
 RA Bosak S.A., Mesman P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huiyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Whiting J., Helton E., Ketterman M., Madan A., Rodriguez S., Sanchez A.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -1- FUNCTION: Component of tight junction (TJ) strands.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: Belongs to the claudin family.
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; AF087825; AA09760.1; -;
 DR EMBL; BC008104; AA08104.1; -;
 DR EMBL; BC050007; AA05007.1; -;
 DR MGI; MGI:1859285; Clad7.
 DR InterPro; IPR006187; Claudin.
 DR InterPro; IPR006188; Claudin_reg.
 DR InterPro; IPR004031; PMP22_Claudin.
 DR Pfam; PF00822; PMP22_Claudin; 1.
 DR PRINTS; PR01077; CLAUDIN.
 DR PROSITE; PS01346; CLAUDIN; 1.
 KM Tight junction; Transmembrane.
 FT TRANSMEM 8 28 POTENTIAL.
 FT TRANSMEM 82 102 POTENTIAL.
 FT TRANSMEM 120 140 POTENTIAL.
 FT TRANSMEM 161 181 POTENTIAL.
 SQ SEQUENCE 211 AA; 22359 MW; 4FE87F3A57AC9F29 CRC64;

Query Match 92.2%; Score 1023; DB 1; Length 211;
 Best Local Similarity 91.9%; Pred. No. 2.2e-81;
 Matches 194; Conservative 7; Mismatches 10; Indels 0; Gaps 0;

QY 1 MANMGQLIGFSGMALGIVGACVTAIPQMOMSSYAGDNIITAOQMYKGLMDCVTOSTG 60
 DB 1 MANSGLQLIGFSGMALGIVGACVTAIPQMOMSSYAGDNIITAOQMYKGLMDCVTOSTG 60
 QY 61 MNSCKMYDSVLAISALQATRALMNVSVLGLAFVATMGKCTRCGGDDVYKARIAM 120
 DB 61 MNSCKMYDSVLAISALQATRALMNVSVLGLAFVATMGKCTRCGGDDVYKARIAM 120
 QY 121 GGGIIFIVAGLAALVACSWYGHQIVTDFYNPILPTNIKEFGPAIFIGAGSALVILGGA 180
 DB 121 TGGIIVIVAGLAALVACSWYGHQIVTDFYNPILPTNIKEFGPAIFIGAGSALVILGGA 180
 QY 181 LLSGSCPGNESKAGRAPRSYKSNSSKEYV 211
 DB 181 LLSGSCPGNESKAGRAPRSYKSNSSKEYV 211

RESULT 3
 CLD7_RAT STANDARD; PRT; 191 AA.
 AC Q92111;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Claudin-7 (Fragment).
 GN CLDN7.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 25, 2004, 02:17:54 ; Search time 232 Seconds

(without alignments)
286,958 Million cell updates/sec

Title: US-10-068-486a-2

Perfect score: 1109
Sequence: 1 MAMMGQLGLRSMALLGWMG.....KAGYRAPRYKSNSEKEY 211

Scoring table:

Gapop 10.0, Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database:

1: SPREMBL_25:*
2: sp_archaea:*
3: sp_bacteria:*
4: sp_fungi:*
5: sp_human:*
6: sp_invertebrate:*
7: sp_mammal:*
8: sp_mhc:*
9: sp_organelle:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriophage:*
17: sp_archaea:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	999	90.1	199	4 Q7Z4Y7	Q7Z4Y7 homo sapien
2	760	68.5	213	13 Q7ZS22	Q7ZS22 xenopus lae
3	701	63.2	211	13 Q8AVG4	Q8AVG4 xenopus lae
4	693	62.5	211	4 Q7Z4X9	Q7Z4X9 homo sapien
5	560	50.5	213	13 Q805G0	Q805G0 xenopus lae
6	555.5	50.1	214	13 Q9DP12	Q9DP12 xenopus lae
7	552.5	49.8	210	13 Q90X08	Q90X08 brachydanio
8	548	49.4	214	13 Q98SR2	Q98SR2 gallus galli
9	547.5	49.4	214	13 Q90XQ9	Q90XQ9 xenopus lae
10	540	48.7	214	13 Q7ZP4	Q7ZP4 brachydanio
11	539.5	48.6	218	13 Q90XR2	Q90XR2 brachydanio
12	538	48.5	214	13 Q90XR0	Q90XR0 brachydanio
13	533	48.1	215	13 Q90XR8	Q90XR8 brachydanio
14	514.5	46.4	206	13 Q90XS1	Q90XS1 brachydanio
15	512.5	46.2	219	11 Q810I9	Q810I9 mus musculus
16	504.5	45.5	266	4 Q8WDM3	Q8WDM3 homo sapien

17	488	44.0	216	13 Q98SR1	Q98SR1 gallus galli
18	470	42.4	209	13 Q90XR9	Q90XR9 brachydanio
19	429	38.7	211	13 Q90XR6	Q90XR6 brachydanio
20	410.5	37.0	224	11 Q8BXA6	Q8BXA6 mus musculus
21	409.5	36.9	210	13 Q90XR4	Q90XR4 brachydanio
22	398	35.9	199	13 Q7ZTS2	Q7ZTS2 brachydanio
23	394.5	35.6	222	13 Q7ZTS7	Q7ZTS7 brachydanio
24	392.5	35.4	237	13 Q90XS0	Q90XS0 brachydanio
25	382	34.4	226	4 Q96N78	Q96N78 homo sapien
26	380	34.3	236	13 Q80HA3	Q80HA3 brachydanio
27	375.5	33.9	236	13 Q7T021	Q7T021 brachydanio
28	373.5	33.7	140	6 Q7YS66	Q7YS66 sus scrofa
29	369.5	33.3	264	11 Q8BZS5	Q8BZS5 mus musculus
30	366.5	33.0	219	13 Q90XR7	Q90XR7 brachydanio
31	361.5	32.6	229	11 Q9CX57	Q9CX57 mus musculus
32	347	31.3	100	6 Q7YS67	Q7YS67 sus scrofa
33	324.5	29.3	194	13 Q7T018	Q7T018 brachydanio
34	308.5	27.8	224	5 Q9N9M1	Q9N9M1 halocynthia
35	307.5	27.7	224	5 Q9N9M2	Q9N9M2 halocynthia
36	289	26.1	193	11 Q8VC62	Q8VC62 mus musculus
37	280.5	25.3	178	4 Q8N7P3	Q8N7P3 homo sapien
38	280.5	25.3	220	11 Q9D7U6	Q9D7U6 mus musculus
39	254.5	22.9	256	13 Q90XR5	Q90XR5 brachydanio
40	227	20.5	159	13 Q8QFL7	Q8QFL7 brachydanio
41	226	20.4	235	6 Q9TUF7	Q9TUF7 bos taurus
42	188.5	17.0	185	13 Q7T020	Q7T020 brachydanio
43	185.5	16.7	296	11 Q9D7D7	Q9D7D7 mus musculus
44	163.5	14.7	111	11 Q924Y3	Q924Y3 rattus norv
45	161	14.5	268	4 Q96B33	Q96B33 homo sapien

ALIGNMENTS

RESULT 1	Q7Z4Y7	PRELIMINARY:	PRT:	199 AA.
AC	Q7Z4Y7			
DT	01-OCT-2003 (TREMBLrel. 25, Created)			
DT	01-OCT-2003 (TREMBLrel. 25, Last sequence update)			
DT	01-OCT-2003 (TREMBLrel. 25, Last annotation update)			
DE	Claudin-1.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_Taxid=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Yue P., Yu L., Bi A.D., Zhang M., He H., Zhao S.Y.;			
RT	"Cloning and expression of a novel human cDNA homology to murine			
RT	Claudin-1 mRNA."			
RL	Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RA	Xu Z.G., Yu L., Yue P., Tu Q., Zheng L.H., Zhao S.Y.;			
RL	Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AF093823; AAP97219.1; -			
SQ	SEQUENCE 199 AA; 21191 MW; E6DABD012C8529C6 CRC64;			
Query Match	90.1%; Score 999; DB 4; Length 199;			
Best Local Similarity	97.0%; Pred. No. 1.9e-80;			
Matches	193; Conservative 0; Mismatches 6; Indels 0; Gaps 0;			
QY	13 MALLGVGLVACTAIPOWMSYVAGDNTTQAAYKGLIMDCVTQSTGMSCKRYDSYLA 72			
DB	1 MALLGGCVSWLCTAIPOWMSYVAGDNTTQAAYKGLIMDCVTQSTGMSCKRYDSYLA 60			
QY	73 ISALIQATRALMVSIVLGFAMFVATMGKCTCGGDDKXKARIMAGGIIITIVAGLA 132			
DB	61 ISALIQATRALMVSIVLGFAMFVATMGKCTCGGDDKXKARIMAGGIIITIVAGLA 120			
QY	133 ALVACSWYGHQIVDFNPILPTNIKEFGAIFIGAGSLVLTGALLSCSGPNESEK 192			

DB 121 ALVACSWYGHQIVTDFYNPLIPNIXEFGPAIFGWAGSALVILGALLSCSCPGNSK 180
 QY 193 AGYRAPRSYPKNSKEYV 211
 DB 181 AGYRAPRSYPKNSKEYV 199

RESULT 2

Q7ZS22 PRELIMINARY; PRT; 213 AA.

AC 07ZS22
 DT 01-JUN-2003 (Tremblrel. 24, Created)
 DT 01-JUN-2003 (Tremblrel. 24, Last sequence update)
 DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
 DE Claudin7L1 (Hypothetical protein).
 GN CLDN7L1.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodinae; Xenopus.
 NC NCB1_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Fujita M., Itoh M., Shibata M., Taira S., Taira M.;
 RT "Gene expression pattern analysis of the tight junction protein,
 Claudin, in the early morphogenesis of Xenopus embryos."
 RL Gene Expr. Patterns 2:23-26(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RA Klein S., Strausberg R.;
 RL Submitted (Mar-2003) to the EMBL/GenBank/DBJ databases.
 DB EMBL; AB072910; BAC21015.1; -
 DB EMBL; BC048771; AAH48771.1; -
 DB GO; GO:0016020; C:membrane; IEA.
 DB GO; GO:0005923; C:tight junction; IEA.
 DB GO; GO:0005198; F:structural molecule activity; IEA.
 DR InterPro: IPR006187; Claudin.reg.
 DR InterPro: IPR006188; Claudin.
 DR InterPro: IPR004031; PMP22_Claudin.
 DR Pfam; PF00822; PMP22_Claudin; 1.
 DR PRINTS; PR01077; CLAUDIN.
 DR PROSITE; PS01346; CLAUDIN; 1.
 DR Hypothetical protein.
 SQ SEQUENCE 213 AA; 22731 MW; CDDF3E2D8C4B5775 CRC64;

Query Match 68.5%; Score 760; DB 13; Length 213;
 Best Local Similarity 65.1%; Pred. No. 2,6e-59;

Matches 140; Conservative 33; Mismatches 36; Indels 6; Gaps 3;

QY 1 MANNGIQLIGFPMALLGVNGLVACTAIFQWQMSYAGNDIITAOQMYKGLMNDVCTOSTG 60
 DT 01-OCT-2003 (Tremblrel. 25, Last sequence update)
 DE Claudin-2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 NC NCB1_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Yang J., Yu L., Zhou Y., Dai F.Y., Xin Y.R., Zhao S.Y.;
 RT "Cloning and sequencing of a new human cDNA homologous to murine
 Claudin-1 mRNA."
 RL Submitted (Jul-2003) to the EMBL/GenBank/DBJ databases.
 DB 121 GGGIIFIVAGLALVACSWYGHQIVTDFYNPLIPNIXEFGPAIFGWAGSALVILGGA 180
 DT 01-OCT-2003 (Tremblrel. 25, Last sequence update)
 DE Claudin-2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 NC NCB1_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Yang J., Yu L., Zhou Y., Dai F.Y., Xin Y.R., Zhao S.Y.;
 RT "Cloning and sequencing of a new human cDNA homologous to murine
 Claudin-1 mRNA."
 RL Submitted (Jul-2003) to the EMBL/GenBank/DBJ databases.
 DB 121 TGGFVFLIGLALALVACSWYGHQIVTDFYNPLIPNIXEFGPAIFGWAGSALVILGGA 180
 DT 01-OCT-2003 (Tremblrel. 25, Last sequence update)
 DE Claudin-2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 NC NCB1_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Yang J., Yu L., Zhou Y., Dai F.Y., Xin Y.R., Zhao S.Y.;
 RT "Cloning and sequencing of a new human cDNA homologous to murine
 Claudin-1 mRNA."
 RL Submitted (Jul-2003) to the EMBL/GenBank/DBJ databases.
 DB 181 LLSGSGCP-GNESKAGYRAPRSYPKS--NSKEYV 211
 DT 01-OCT-2003 (Tremblrel. 25, Last sequence update)
 DE Claudin-2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 NC NCB1_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Yang J., Yu L., Zhou Y., Dai F.Y., Xin Y.R., Zhao S.Y.;
 RT "Cloning and sequencing of a new human cDNA homologous to murine
 Claudin-1 mRNA."
 RL Submitted (Jul-2003) to the EMBL/GenBank/DBJ databases.
 DB 181 LLSGSGCRNNYQKGY--FKSGAKSKVPSGRDYV 213
 DT 01-OCT-2003 (Tremblrel. 25, Last sequence update)
 DE Claudin-2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 NC NCB1_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Yang J., Yu L., Zhou Y., Dai F.Y., Xin Y.R., Zhao S.Y.;
 RT "Cloning and sequencing of a new human cDNA homologous to murine
 Claudin-1 mRNA."
 RL Submitted (Jul-2003) to the EMBL/GenBank/DBJ databases.

RESULT 3

Q8AVG4 PRELIMINARY; PRT; 211 AA.

AC 08AVG4
 DT 01-MAR-2003 (Tremblrel. 23, Created)
 DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)

DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)

DE Similar to claudin 1.

OS Xenopus laevis (African clawed frog).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;

OC Xenopodinae; Xenopus.

NC NCB1_TaxID=8355;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Embryo;

RA Klein S., Strausberg R.;

RL Submitted (Jan-2003) to the EMBL/GenBank/DBJ databases.

DB EMBL; BC042293; AAH42293.1; -

DB GO; GO:0016020; C:membrane; IEA.

DB GO; GO:0005923; C:tight junction; IEA.

DB GO; GO:0005198; F:structural molecule activity; IEA.

DR InterPro: IPR006187; Claudin.reg.

DR InterPro: IPR006188; Claudin.

DR Pfam; PF00822; PMP22_Claudin.

DR PRINTS; PR01077; CLAUDIN.

DR PROSITE; PS01346; CLAUDIN; 1.

SQ SEQUENCE 211 AA; 22476 MW; 89BA9C7C36F28C98 CRC64;

Query Match 63.2%; Score 701; DB 13; Length 211;
 Best Local Similarity 60.6%; Pred. No. 4,1e-54;

Matches 129; Conservative 31; Mismatches 49; Indels 4; Gaps 2;

QY 1 MANNGIQLIGFPMALLGVNGLVACTAIFQWQMSYAGNDIITAOQMYKGLMNDVCTOSTG 60
 DT 01-OCT-2003 (Tremblrel. 25, Last sequence update)
 DE Claudin-2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 NC NCB1_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Yang J., Yu L., Zhou Y., Dai F.Y., Xin Y.R., Zhao S.Y.;
 RT "Cloning and sequencing of a new human cDNA homologous to murine
 Claudin-1 mRNA."
 RL Submitted (Jul-2003) to the EMBL/GenBank/DBJ databases.
 DB 121 GGGIIFIVAGLALVACSWYGHQIVTDFYNPLIPNIXEFGPAIFGWAGSALVILGGA 180
 DT 01-OCT-2003 (Tremblrel. 25, Last sequence update)
 DE Claudin-2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 NC NCB1_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Yang J., Yu L., Zhou Y., Dai F.Y., Xin Y.R., Zhao S.Y.;
 RT "Cloning and sequencing of a new human cDNA homologous to murine
 Claudin-1 mRNA."
 RL Submitted (Jul-2003) to the EMBL/GenBank/DBJ databases.
 DB 121 VGGALFIVAGLCTVLTATWYGDKXKADFNMFTPTNSKTEFGPAIFGWAGSALVILGGA 180
 DT 01-OCT-2003 (Tremblrel. 25, Last sequence update)
 DE Claudin-2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 NC NCB1_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Yang J., Yu L., Zhou Y., Dai F.Y., Xin Y.R., Zhao S.Y.;
 RT "Cloning and sequencing of a new human cDNA homologous to murine
 Claudin-1 mRNA."
 RL Submitted (Jul-2003) to the EMBL/GenBank/DBJ databases.
 DB 181 LLSGSGCPNSKAGYRAPRSYPKS--NSKEYV 211
 DT 01-OCT-2003 (Tremblrel. 25, Last sequence update)
 DE Claudin-2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 NC NCB1_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Yang J., Yu L., Zhou Y., Dai F.Y., Xin Y.R., Zhao S.Y.;
 RT "Cloning and sequencing of a new human cDNA homologous to murine
 Claudin-1 mRNA."
 RL Submitted (Jul-2003) to the EMBL/GenBank/DBJ databases.
 DB 181 LLSGSGCRNNYQKGY--FKSGAKSKVPSGRDYV 211
 DT 01-OCT-2003 (Tremblrel. 25, Last sequence update)
 DE Claudin-2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 NC NCB1_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Yang J., Yu L., Zhou Y., Dai F.Y., Xin Y.R., Zhao S.Y.;
 RT "Cloning and sequencing of a new human cDNA homologous to murine
 Claudin-1 mRNA."
 RL Submitted (Jul-2003) to the EMBL/GenBank/DBJ databases.

Query Match 62.5%; Score 693; DB 4; Length 211;
 Best Local Similarity 59.6%; Pred. No. 2,1e-53;

Matches 127; Conservative 38; Mismatches 44; Indels 4; Gaps 2;